

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 6, 2006, 09:22:55 ; Search time 78.2895 Seconds
(without alignments)
39.286 Million cell updates/sec

Title: US-10-632-388-296

Perfect score: 37

Sequence: 1 ILAPPVP 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Database :

A_Geneseq_21.*
1: Geneseqp1980s.*
2: Geneseqp1990s.*
3: Geneseqp2000s.*
4: Geneseqp2001s.*
5: Geneseqp2002s.*
6: Geneseqp2003as.*
7: Geneseqp2003bs.*
8: Geneseqp2004s.*
9: Geneseqp2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	37	100.0	7	3	AAB17240 SH3 antag
2	37	100.0	7	5	ABB73233 Src homol
3	37	100.0	7	7	ADJ73387 SH3 antag
4	37	100.0	7	8	ADJ53021 CH1 delet
5	37	100.0	7	8	ADJ51982 CH1 delet
6	37	100.0	13	2	AAW05414 Src SH3 d
7	37	100.0	13	2	AAW05481 SH3-Bindi
8	37	100.0	13	2	AAW11100 Src SH3 d
9	37	100.0	13	2	AAW11117 Src SH3 d
10	37	100.0	13	2	AAW38066 PPPPY mot
11	37	100.0	13	2	AAW25513 SH3 synth
12	37	100.0	13	7	ADB49256 Biotinyia
13	37	100.0	16	2	AAW24028 Synthetic
14	37	100.0	17	2	AAW05412 Src SH3 d
15	37	100.0	30	2	AAW16938 Random re
16	37	100.0	30	2	AAW25501 Random pe
17	37	100.0	1767	4	ABB62726 Drosophil
18	36	97.3	16	2	ABP25387 PLC gamma
19	36	97.3	137	5	ABP43530 Human sec
20	36	97.3	159	4	AAG83323 P patens
21	36	97.3	159	4	AAG80875 Lipid bio
22	36	97.3	200	4	ABB50640 Human sec
23	36	97.3	200	6	ABO44897 Novel hum
24	36	97.3	200	7	ABO26377 Protein a

25	36	97.3	218	8	ADQ67709
26	35	94.6	116	4	AAU17916
27	35	94.6	116	7	ADG41296
28	35	94.6	116	7	ADI97070
29	35	94.6	132	3	AAI10732
30	35	94.6	477	7	ABM90479
31	35	94.6	477	7	ABM87558
32	34	91.9	113	7	ADB65024
33	34	91.9	177	7	ADC97354
34	34	91.9	291	5	ADK34392
35	33	89.2	38	9	ADV58105
36	33	89.2	39	9	ADV59047
37	33	89.2	67	7	ADC94833
38	33	89.2	93	2	AAV00298
39	33	89.2	93	7	ADD90327
40	33	89.2	93	7	ADG90146
41	33	89.2	93	9	ADY25486
42	33	89.2	95	7	ADP60277
43	33	89.2	112	4	ABB17509
44	33	89.2	161	7	ABO68453
45	33	89.2	190	4	ABG27197
46	33	89.2	191	8	ADX74877
47	33	89.2	198	7	ADC31836
48	33	89.2	207	7	ADD26860
49	33	89.2	267	4	AAU14624
50	33	89.2	285	8	ADX77177
51	33	89.2	287	7	ABO80729
52	33	89.2	322	7	ABO76531
53	33	89.2	342	7	ADF59220
54	33	89.2	363	6	ABU19760
55	33	89.2	439	8	ADH71374
56	33	89.2	443	5	AAU91280
57	33	89.2	478	8	ADU46858
58	33	89.2	478	8	ADO32024
59	33	89.2	547	2	AAV14566
60	33	89.2	551	2	AAV14567
61	33	89.2	629	8	ADH71372
62	33	89.2	637	6	AAO26517
63	33	89.2	695	8	ADP05691
64	33	89.2	705	9	ABM92061
65	33	89.2	742	8	ADH71368
66	33	89.2	841	7	ABM87673
67	33	89.2	984	1	ABP80927
68	33	89.2	984	2	AAV21622
69	33	89.2	984	5	ABP54693
70	33	89.2	984	8	ADL12398
71	33	89.2	984	8	ADN04124
72	33	89.2	984	8	ADP05689
73	33	89.2	984	8	ADQ80383
74	33	89.2	1060	6	ABP81961
75	33	89.2	1243	5	ABP69341
76	33	89.2	1248	6	ADA37508
77	33	89.2	1248	8	ADE81083
78	33	89.2	1278	6	ADA37510
79	33	89.2	1278	8	ADE81085
80	33	89.2	1307	4	ABE65464
81	33	89.2	1329	5	ABE90782
82	33	89.2	1329	5	ABE90728
83	33	89.2	1329	6	ABU54489
84	33	89.2	1329	6	ABU54435
85	33	89.2	1329	8	ADO29669
86	33	89.2	1330	8	ADH71364
87	33	89.2	1331	5	AAU74824
88	33	89.2	1331	5	ABB90725
89	33	89.2	1331	5	ABB90740
90	33	89.2	1331	6	ABU54432
91	33	89.2	1331	6	ABU54447
92	33	89.2	1331	8	ADO29668
93	33	89.2	1338	4	AAU71863
94	33	89.2	1338	7	AAO30115
95	33	89.2	1367	8	ADN22855
96	33	89.2	1367	8	ADN22856
97	33	89.2	1512	7	ADE08730

ADG67709	Novel hum
AAU17916	Novel hum
ADG41296	Human res
ADI97070	Human res
AAI10732	Arabidops
ABM90479	Rice abio
ABM87558	Rice abio
ADB65024	Human pro
ADC97354	E. faeciu
ADK34392	Novel hum
ADV58105	G protein
ADV59047	G protein
ADC94833	E. faeciu
AAV00298	Human sec
ADD90327	Novel hum
ADG90146	Human pro
ADY25486	Novel hum
ADF60277	Human con
ABB17509	Human ner
ABO68453	Pseudomon
ABG27197	Novel hum
ADX74877	Plant ful
ADC31836	Human nov
ADD26860	Human adi
AAU14624	Novel bon
ADX77177	Plant ful
ABO80729	Pseudomon
ABO76531	Pseudomon
ADF59220	Human pol
ABU19760	Protein e
ADH71374	Human pro
AAU91280	Human NOV
ADJ46858	Human KIA
ADO32024	Cyclin-de
AAV14566	Env prote
AAV14567	Alternati
ADH71372	Human pro
AAO26517	Protein o
ADP05691	Mouse nuc
ABM92061	M. xanthu
ADH71368	Human pro
ABM87673	Rice abio
ABP80927	Sequence
AAV21622	Ligand bi
ABP54693	Metastati
ADL12398	Human ste
ADN04124	Antipsoxi
ADP05689	Human nuc
ADQ80383	Nuclear r
ABP81961	Human G p
ABP69341	Human pol
ADA37508	Human G-p
ADE81083	Novel G p
ADA37510	Human G-p
ADE81085	Novel G p
ABE65464	Drosophil
ABE90782	Mouse Tum
ABE90728	Mouse Tum
ABU54489	Mouse tum
ABU54435	Mouse tum
ADO29669	Mouse GPC
ADH71364	Human pro
AAU74824	Human REP
ABB90725	Human Tum
ABB90740	Human Tum
ABU54432	Human tum
ABU54447	Human tum
ADO29668	Human GPC
AAU71863	Human h15
Aao30115	Human tum
ADN22855	Bacterial
ADN22856	Bacterial
ADE08730	Novel pro

98	33	89.2	1542	8	ADH71384	Adh71384 Human pro	171	32	86.5	1299	4	AAW78454	Aam78454 Human pro
99	33	89.2	1542	8	ADH71362	Adh71362 Human pro	172	32	86.5	1299	9	ADZ09816	Adz09816 Human bre
100	33	89.2	1542	8	ADH71370	Adh71370 Human pro	173	32	86.5	1307	4	AAW79438	Aam79438 Human pro
101	33	89.2	6310	6	ABU39869	Abu39869 Protein e	174	32	86.5	1307	4	AAW79439	Aam79439 Human pro
102	32	86.5	16	2	AAW25421	Aaw25421 Crk N-ter	175	32	86.5	1329	9	ADX07105	Abx07105 Cyclin-de
103	32	86.5	22	4	AAW16556	Aam16556 Peptide #	176	32	86.5	1630	4	ABW44550	Abw44550 Human wou
104	32	86.5	22	4	ABR35540	Abb35540 Peptide #	177	32	86.5	1630	6	ABP98331	Abp98331 Amino aci
105	32	86.5	22	4	AAW29038	Aam29038 Peptide #	178	32	86.5	1632	4	AAW39023	Aam39023 Human pol
106	32	86.5	22	4	ABR30366	Abb30366 Peptide #	179	32	86.5	2171	7	ADC99058	Adc99058 Human rpp
107	32	86.5	22	4	ABR20964	Abb20964 Protein #	180	32	86.5	2955	4	ABW21509	Abw21509 Novel hum
108	32	86.5	22	4	AAW68729	Aam68729 Human bon	181	32	86.5	2958	4	ABW44617	Abw44617 Human wou
109	32	86.5	22	4	AAW56353	Aam56353 Human bra	182	32	86.5	3530	8	ADQ18982	Adq18982 Human eof
110	32	86.5	22	4	ABG50398	Abg50398 Human liv	183	32	86.5	31267	6	ABG74786	Abg74786 Human rgs
111	32	86.5	22	4	AAW04269	Aam04269 Peptide #	184	32	86.5	34350	8	ADQ89964	Adq89964 Antagonis
112	32	86.5	22	5	ABG38314	Abg38314 Human pep	185	31	83.8	16	2	AAW25334	Aaw25334 Cortactin
113	32	86.5	28	3	AAW64765	Aay64765 Human 5'	186	31	83.8	20	2	AAW16957	Aaw16957 Cortactin
114	32	86.5	28	9	ADU72329	Adu72329 Signal pe	187	31	83.8	97	8	AAO21466	Aao21466 Eukaryoti
115	32	86.5	28	9	ADU73320	Adu73320 Human inc	188	31	83.8	98	8	ADU73235	Adu73235 Plant ful
116	32	86.5	61	4	AAU60444	Aau60444 Propionib	189	31	83.8	108	4	ABG23885	Abg23885 Novel hum
117	32	86.5	61	6	ABW56863	Abm56863 Propionib	190	31	83.8	121	4	AAU46725	Aau46725 Propionib
118	32	86.5	80	8	ADU79019	Adu79019 Plant ful	191	31	83.8	121	6	ABM43244	Abm43244 Propionib
119	32	86.5	84	5	ABP06885	Abp06885 Human ORF	192	31	83.8	122	4	AAO09541	Aao09541 Novel hum
120	32	86.5	101	5	ABP26410	Abp26410 Streptoco	193	31	83.8	132	5	ADK34787	Adk34787 Novel hum
121	32	86.5	101	8	ADV81495	Adv81495 Streptoco	194	31	83.8	143	7	ABO69656	Abob69656 Pseudom
122	32	86.5	143	4	ABE71001	Abv71001 Drosophi	195	31	83.8	174	9	ABM94517	Abm94517 M. xanthu
123	32	86.5	151	7	ADE72523	Ade72523 Human end	196	31	83.8	177	7	ABO71055	Abob71055 Pseudom
124	32	86.5	151	7	ADE72522	Ade72522 Human end	197	31	83.8	189	9	ABE39379	Abob39379 L. pneumo
125	32	86.5	151	8	ADU88060	Adu88060 Streptoco	198	31	83.8	190	2	AAW88004	Aaw88004 Protein e
126	32	86.5	151	8	ADU79313	Adu79313 Streptoco	199	31	83.8	191	9	ABE35945	Abob35945 L. pneumo
127	32	86.5	164	4	AAU42376	Aau42376 Propionib	200	31	83.8	269	8	ADS30919	Ads30919 Bacteri
128	32	86.5	164	6	ABM38895	Abm38895 Propionib	201	31	83.8	274	4	AAU42242	Aau42242 Propionib
129	32	86.5	217	9	ABM95613	Abm95613 M. xanthu	202	31	83.8	274	4	AAU42242	Aau42242 Propionib
130	32	86.5	282	3	AGU5722	Agau5722 Arabidops	203	31	83.8	302	6	ABM38761	Abm38761 Propionib
131	32	86.5	297	3	AGU5721	Agau5721 Arabidops	204	31	83.8	310	2	AAU31953	Agau31953 Plant ful
132	32	86.5	315	3	ADC87269	Adc87269 Human GPC	205	31	83.8	310	7	ADK68446	Adk68446 Simlan T-
133	32	86.5	331	3	AGU50720	Agau50720 Arabidops	206	31	83.8	310	8	ADU49881	Adu49881 STLV-3 en
134	32	86.5	332	9	ADY30823	Ady30823 Thale cre	207	31	83.8	325	8	ADU68241	Adu68241 Plant ful
135	32	86.5	419	2	AAU50558	Aau50558 HSV2 LAT	208	31	83.8	332	3	AAU55461	Agau55461 Arabidops
136	32	86.5	419	7	ADG75126	Adg75126 Human her	209	31	83.8	341	3	AAU55460	Agau55460 Arabidops
137	32	86.5	433	5	AAU50962	AAU50962 Maize met	210	31	83.8	347	2	AAW88005	AAW88005 GA4 homol
138	32	86.5	433	6	ABU00021	Abu00021 Maize met	211	31	83.8	347	5	ABB91678	Abb91678 Herbicida
139	32	86.5	433	6	ABU00021	Abu00021 Human nov	212	31	83.8	379	2	AAU31954	Agau31954 Glutathio
140	32	86.5	433	7	ADG19014	Adg19014 Maize 2mM	213	31	83.8	381	8	ADT78207	Adt78207 Plant ful
141	32	86.5	453	8	ADY06683	Ady06683 Plant ful	214	31	83.8	387	8	ADT58908	Adt58908 Plant pol
142	32	86.5	460	4	ABE71622	Abv71622 Drosophi	215	31	83.8	406	4	AAU95429	Agau95429 Human pro
143	32	86.5	465	6	ABM42598	Abm42598 Propionib	216	31	83.8	406	4	AAU32015	Agau32015 Novel hum
144	32	86.5	465	6	ABM42598	Abm42598 Propionib	217	31	83.8	445	4	AAU94676	Agau94676 Human pro
145	32	86.5	611	9	ADY30825	Ady30825 Thale cre	218	31	83.8	445	7	ADC31765	Adc31765 Human nov
146	32	86.5	611	9	ADY30825	Ady30825 Thale cre	219	31	83.8	445	8	ABM81200	Abm81200 Tumour-as
147	32	86.5	613	4	ABR94287	Abb94287 Human pro	220	31	83.8	457	4	ABG02182	Abg02182 Novel hum
148	32	86.5	633	6	AAO16745	Aao16745 Aspergill	221	31	83.8	481	4	ABG63482	Abg63482 Drosophi
149	32	86.5	638	3	AGU38502	Agau38502 Arabidops	222	31	83.8	486	7	ADU47729	Adu47729 Human pro
150	32	86.5	659	7	ADC32976	Adc32976 Human nov	223	31	83.8	492	2	AAU85471	Agau85471 PTLV-L en
151	32	86.5	683	8	ADT55490	Adt55490 Plant pol	224	31	83.8	522	4	ABB70785	Abb70785 Drosophi
152	32	86.5	685	7	ADP64762	Adp64762 Human pro	225	31	83.8	524	6	ABU22998	Abu22998 Protein e
153	32	86.5	694	5	ABF63094	Abf63094 FLO11 gen	226	31	83.8	593	7	ABM89613	Abm89613 Rice abio
154	32	86.5	720	3	ABG38501	Agau38501 Arabidops	227	31	83.8	614	7	ABO71464	Abob71464 Pseudom
155	32	86.5	740	6	ABJ26424	Abj26424 Aspergill	228	31	83.8	854	9	ABG16509	Abg16509 Novel hum
156	32	86.5	740	6	ABJ26424	Abj26424 Aspergill	229	31	83.8	854	9	ADX39652	Adx39652 HIV Env p
157	32	86.5	777	3	AGU38500	Agau38500 Arabidops	230	31	83.8	854	9	ADX39659	Adx39659 HIV Env p
158	32	86.5	779	8	ADX95214	Adx95214 Plant ful	231	31	83.8	856	9	ADX39650	Adx39650 HIV Env p
159	32	86.5	808	5	ABR505653	Abb505653 Human DNA	232	31	83.8	872	9	ADX39755	Adx39755 HIV Env p
160	32	86.5	885	8	ADP22490	Adp22490 Sea-squir	233	31	83.8	873	7	ADU07846	Adu07846 Rice prot
161	32	86.5	998	4	ABU53140	Abu53140 Human tes	234	31	83.8	949	4	ABE59399	Abbe59399 Drosophi
162	32	86.5	1024	9	ADC31296	Adc31296 Human nov	235	31	83.8	1016	8	ADR08525	Adr08525 Human pro
163	32	86.5	1048	9	ADZ70351	Adz70351 Human pro	236	31	83.8	1027	7	ADJ69375	Adj69375 Human hpa
164	32	86.5	1083	9	ADZ80792	Adz80792 Amino aci	237	31	83.8	1031	7	ABM88370	Abm88370 Rice abio
165	32	86.5	1100	4	AAE933286	Aae933286 Human pro	238	31	83.8	1085	2	AAU90121	Aau90121 Human HUB
166	32	86.5	1100	6	AAO29563	Aao29563 Human Pci	239	31	83.8	1085	2	AAU41646	Aau41646 Human BUB
167	32	86.5	1102	8	ADY70491	Ady70491 Human bet	240	31	83.8	1085	3	AAU59142	Aau59142 Human eer
168	32	86.5	1192	8	ADN20849	Adn20849 Bacteri	241	31	83.8	1085	5	AAE25908	Aae25908 Human BUB
169	32	86.5	1258	4	AAW78455	Aam78455 Human pro	242	31	83.8	1085	6	ABG75608	Abg75608 Human BUB
170	32	86.5	1299	3	AAU58633	Aau58633 Protein r	243	31	83.8	1085	7	ADG18109	Adg18109 Human kin

244	31	83.8	1085	7	ADG18095	Adg18095 Human kin	317	30	81.1	125	4	AAm68679	AAm68679 Human bon
245	31	83.8	1085	7	ADN95253	Adn95253 Human BEC	318	30	81.1	125	4	AAm56301	AAm56301 Human bra
246	31	83.8	1085	8	ADK67722	Adk67722 Human mod	319	30	81.1	125	4	ABG50343	ABG50343 Human liv
247	31	83.8	1085	8	ADO00999	Ado00999 Human hom	320	30	81.1	125	4	AAm04220	AAm04220 Peptide #
248	31	83.8	1085	8	ADO19728	Ado19728 Human PRO	321	30	81.1	125	5	ABG38258	ABG38258 Human pep
249	31	83.8	1085	8	ADQ09188	Adq09188 Human BUB	322	30	81.1	127	3	AAg26180	AAg26180 Zea may
250	31	83.8	1085	9	ADY16259	Ady16259 PRO polyP	323	30	81.1	128	7	ADL06668	ADL06668 Human 3T3
251	31	83.8	1095	2	AAy06287	Aay06287 Human BUB	324	30	81.1	128	7	ADM05445	ADM05445 Human pro
252	31	83.8	1141	8	ADK60415	Adk60415 Angiogene	325	30	81.1	132	4	AAO07484	AAO07484 Human pol
253	31	83.8	1141	8	ADK60716	Adk60716 Angiogene	326	30	81.1	137	8	ADY08238	ADY08238 Plant ful
254	31	83.8	1141	8	ADP73339	Adp73339 Human mit	327	30	81.1	139	3	AAg23324	AAg23324 Arabidops
255	31	83.8	1175	4	ABG17468	Abg17468 Novel hum	328	30	81.1	139	3	AAg36753	AAg36753 Arabidops
256	31	83.8	1175	5	Aau76512	Aau76512 Human HPK	329	30	81.1	141	3	AAg23323	AAg23323 Arabidops
257	31	83.8	1175	8	ADK60186	Adk60186 Angiogene	330	30	81.1	142	4	AAU04823	AAU04823 Micromono
258	31	83.8	1175	8	ADK60487	Adk60487 Angiogene	331	30	81.1	144	4	AAm87652	AAm87652 Human imm
259	31	83.8	1175	8	ADP73110	Adp73110 Angiogene	332	30	81.1	145	4	AAU61889	AAU61889 Propionib
260	31	83.8	1212	4	AAE04368	Aae04368 Human kin	333	30	81.1	145	6	ABM58408	ABM58408 Propionib
261	31	83.8	1212	7	ADK25753	Adk25753 Human pro	334	30	81.1	147	3	AAg23322	AAg23322 Arabidops
262	31	83.8	1212	8	ADT98449	Adt98449 Human Mit	335	30	81.1	150	7	ABO79053	ABO79053 Pseudomon
263	31	83.8	1219	4	ABG17469	Abg17469 Novel hum	336	30	81.1	151	2	AAW54438	AAW54438 Mouse nov
264	31	83.8	1230	2	AAW48895	Aaw48895 Candida a	337	30	81.1	151	2	AAy37644	AAy37644 Amino aci
265	31	83.8	1239	2	AAy55931	Aay55931 Human ZC1	338	30	81.1	151	3	AAb10245	AAb10245 Murine ad
266	31	83.8	1239	8	ADT98448	Adt98448 Human Mit	339	30	81.1	151	4	AAO03607	AAO03607 Human pol
267	31	83.8	1239	8	ADT98452	Adt98452 Human ger	340	30	81.1	154	3	AAO07339	AAO07339 Arabidops
268	31	83.8	1320	8	ADT98450	Adt98450 Human Mit	341	30	81.1	158	3	ABa41373	ABa41373 Human ORF
269	31	83.8	1384	8	ADN61482	Adn61482 Human KPP	342	30	81.1	160	8	ADR09710	ADR09710 Human pro
270	31	83.8	1392	8	ADJ96675	Adj96675 Human pro	343	30	81.1	163	4	AAm25236	AAm25236 Human pro
271	31	83.8	1655	4	ABG17466	Abg17466 Novel hum	344	30	81.1	163	4	ABB12270	ABb12270 Human gua
272	31	83.8	1665	4	ABB64010	Abb64010 Drosophil	345	30	81.1	163	5	ABU05941	ABu05941 M. tuberc
273	31	83.8	1726	7	ABM88576	Abm88576 Rice abio	346	30	81.1	163	6	ABM69378	ABm69378 Phototrab
274	31	83.8	2080	7	ABE31679	Abe31679 Human 707	347	30	81.1	164	7	ABO82761	ABO82761 Pseudomon
275	31	83.8	4881	3	AAE23751	Aae23751 S. avermi	348	30	81.1	167	4	ABG19252	ABg19252 Novel hum
276	31	83.8	4881	4	AAE65266	Aae65266 Streptomy	349	30	81.1	171	7	ABM89498	ABm89498 Rice abio
277	31	83.8	10	4	ABB81441	Abb81441 Target cy	350	30	81.1	177	8	ADP56387	ADp56387 Human PRO
278	30	81.1	24	6	ADB49304	Adb49304 Novel WW	351	30	81.1	178	8	ABO59894	ABO59894 Human gen
279	30	81.1	38	6	ABJ19800	Abj19800 DOCK 3 tu	352	30	81.1	198	4	ABO93283	ABO93283 Human pro
280	30	81.1	44	4	AAm88374	Aam88374 Human imm	353	30	81.1	200	8	ADX68196	ADx68196 Plant ful
281	30	81.1	52	4	ABG11384	Abg11384 Novel hum	354	30	81.1	201	8	ADY23460	ADy23460 Plant ful
282	30	81.1	62	4	AAU43820	Aau43820 Propionib	355	30	81.1	202	8	ABM82069	ABm82069 Tumour-as
283	30	81.1	62	6	ABM40339	Abm40339 Propionib	356	30	81.1	203	7	ADF04044	ADF04044 Bacterial
284	30	81.1	64	4	AAU45941	Aau45941 Propionib	357	30	81.1	204	3	AAG54583	AAg54583 Zea may
285	30	81.1	64	6	ABM42460	Abm42460 Propionib	358	30	81.1	208	4	AAE01981	AAe01981 Human ATP
286	30	81.1	68	4	AAO02847	Aao02847 Human pol	359	30	81.1	214	8	ADX71182	ADx71182 Plant ful
287	30	81.1	69	4	AAU40618	Aau40618 Propionib	360	30	81.1	217	8	ADX96879	ADx96879 Plant ful
288	30	81.1	69	6	ABM37137	Abm37137 Propionib	361	30	81.1	218	2	AAW59198	AAw59198 Seq ID 92
289	30	81.1	79	3	AAO02340	Aao02340 Human sec	362	30	81.1	218	2	AAW40087	AAw40087 Seq ID 92
290	30	81.1	81	4	AAU47970	Aau47970 Propionib	363	30	81.1	219	2	AAW59197	AAw59197 Seq ID 91
291	30	81.1	81	6	ABM44489	Abm44489 Propionib	364	30	81.1	219	2	AAW40086	AAw40086 Seq ID 91
292	30	81.1	85	5	ABG62118	Abg62118 Human pro	365	30	81.1	226	7	ADI40526	ADI40526 Human pur
293	30	81.1	90	4	AAU39513	Aau39513 Propionib	366	30	81.1	226	8	ABO59343	ABO59343 Human gen
294	30	81.1	90	6	ABM36032	Abm36032 Propionib	367	30	81.1	231	4	ADC86879	ADC86879 Human GPC
295	30	81.1	93	5	ADK35932	Adk35932 Novel hum	368	30	81.1	232	4	ABG09954	ABg09954 Novel hum
296	30	81.1	98	4	AAO08196	Aao08196 Human pol	369	30	81.1	233	8	ADI42825	ADI42825 Plant tra
297	30	81.1	99	4	AAO07957	Aao07957 Human pol	370	30	81.1	233	8	ADO03021	ADO03021 Thalecree
298	30	81.1	100	2	AAy19684	Aay19684 SEQ ID NO	371	30	81.1	236	4	AAm42101	AAm42101 Human pol
299	30	81.1	101	7	ADM06137	Adm06137 Human pro	372	30	81.1	239	8	ADX88341	ADx88341 Plant ful
300	30	81.1	109	7	ADK42061	Adk42061 Human can	373	30	81.1	242	8	ADX72374	ADx72374 Plant ful
301	30	81.1	109	8	ADL83122	Adl83122 Human PRO	374	30	81.1	244	4	ABB62905	ABb62905 Drosophil
302	30	81.1	109	8	ADO19902	Ado19902 Human PRO	375	30	81.1	246	2	AAr29894	AAr29894 HCV NS4-N
303	30	81.1	109	8	ADP24582	Adp24582 PRO polyP	376	30	81.1	247	8	ADX76687	ADx76687 Plant ful
304	30	81.1	112	4	AAU45409	Aau45409 Propionib	377	30	81.1	250	6	AAO26516	AAO26516 Protein o
305	30	81.1	112	6	ABM41928	Abm41928 Propionib	378	30	81.1	255	3	AAb33114	AAb33114 Pinus rad
306	30	81.1	113	8	ABO58059	AbO58059 Human gen	379	30	81.1	257	3	AAb25424	AAb25424 Pinus rad
307	30	81.1	115	4	AAU50188	Aau50188 Propionib	380	30	81.1	260	7	ABM89772	ABm89772 Rice abio
308	30	81.1	115	5	ABP34174	Abp34174 Human ORF	381	30	81.1	263	5	ABP51376	ABp51376 Human MDP
309	30	81.1	115	6	ABM46707	Abm46707 Propionib	382	30	81.1	264	5	ABP43591	ABp43591 Human cLo
310	30	81.1	122	3	ADG39025	Adg39025 Arabidops	383	30	81.1	264	8	ADT07510	ADt07510 Human col
311	30	81.1	122	8	ADX72660	Adx72660 Plant ful	384	30	81.1	268	7	ADG42188	ADg42188 Human bra
312	30	81.1	125	4	AAW16497	Aaw16497 Peptide #	385	30	81.1	272	7	ABO80555	ABO80555 Pseudomon
313	30	81.1	125	4	ABB35483	Abb35483 Peptide #	386	30	81.1	273	4	ABG24064	ABg24064 Novel hum
314	30	81.1	125	4	AAW28985	Aaw28985 Peptide #	387	30	81.1	273	7	ABO73472	ABO73472 Pseudomon
315	30	81.1	125	4	ABB30311	Abb30311 Peptide #	388	30	81.1	274	8	ADP67314	ADp67314 Human bla
316	30	81.1	125	4	ABB20920	Abb20920 Protein #	389	30	81.1	276	8	ADX92319	ADx92319 Plant ful

390	30	81.1	278	7	ADL07468	Adl07468	Wheat TaD	463	30	81.1	481	7	ABO2916	Pseudomon
391	30	81.1	278	8	ADT59573	Adt59573	Plant pol	464	30	81.1	486	8	ADX96724	Plant ful
392	30	81.1	278	8	ADT76499	Adt76499	Plant ful	465	30	81.1	491	7	ABM85589	Human pro
393	30	81.1	279	8	ADM06052	Adm06052	Human pro	466	30	81.1	495	8	ADY23100	Plant ful
394	30	81.1	279	8	ADY05913	Ady05913	Plant ful	467	30	81.1	497	7	ADD14148	Human hca
395	30	81.1	281	4	ABB63034	Abb63034	Drosophil	468	30	81.1	524	7	ADJ70769	Human era
396	30	81.1	282	4	AAU32765	Aau32765	Novel hum	469	30	81.1	527	2	AAV49513	Xylulokin
397	30	81.1	285	8	ADX95838	Adx95838	Plant ful	470	30	81.1	527	8	ADQ89930	Antagonis
398	30	81.1	286	8	ADX95837	Adx95837	Plant ful	471	30	81.1	527	8	ADR66732	Human pro
399	30	81.1	287	9	AEA20252	Aea20252	Novel hum	472	30	81.1	527	8	ADR66390	Human pro
400	30	81.1	288	8	ADY12581	Ady12581	Plant ful	473	30	81.1	531	9	ADW08753	Human pro
401	30	81.1	288	8	ADY06457	Ady06457	Plant ful	474	30	81.1	534	2	AAW25769	Human MLN
402	30	81.1	293	3	AGG06915	Agg06915	Arabidops	475	30	81.1	534	8	ADR58951	Human Blk
403	30	81.1	293	3	AGG28719	Agg28719	Arabidops	476	30	81.1	537	8	ADQ20678	Human sof
404	30	81.1	293	3	AGG40059	Agg40059	Arabidops	477	30	81.1	540	8	ADT57268	Plant pol
405	30	81.1	294	8	ADG78274	Adg78274	Plant ful	478	30	81.1	541	9	AEA62712	Mitochond
406	30	81.1	294	8	ADG72801	Adg72801	Plant ful	479	30	81.1	553	7	ABO68962	Pseudomon
407	30	81.1	296	3	AGG07338	Agg07338	Arabidops	480	30	81.1	553	7	ADC38726	Human sec
408	30	81.1	296	8	ADN74745	Adn74745	Thale cre	481	30	81.1	568	8	ADQ67693	Novel hum
409	30	81.1	297	9	AEA21066	Aea21066	Novel hum	482	30	81.1	572	2	AAW31855	Mycobacte
410	30	81.1	301	4	AAU36094	Aau36094	Klebsiell	483	30	81.1	575	8	ADT07559	Human col
411	30	81.1	301	6	ABU32009	Abu32009	Protein e	484	30	81.1	577	8	ADT07561	Human col
412	30	81.1	304	3	AGG06914	Agg06914	Arabidops	485	30	81.1	578	4	AAE01985	Human ATP
413	30	81.1	304	3	AGG28718	Agg28718	Arabidops	486	30	81.1	579	8	AAO24552	Murine BH
414	30	81.1	304	3	AGG40058	Agg40058	Arabidops	487	30	81.1	583	6	ABM66070	Propionib
415	30	81.1	304	4	ABM63066	Abm63066	Drosophil	488	30	81.1	583	9	ADW18471	Pinus rad
416	30	81.1	304	7	ABO66897	Abu66897	Plant ful	489	30	81.1	585	8	ADT07557	Human col
417	30	81.1	304	8	ADY09972	Ady09972	Plant ful	490	30	81.1	586	7	ADT07561	Human col
418	30	81.1	305	2	AAW39356	Aaw39356	Novel hum	491	30	81.1	595	6	ABU33902	Protein e
419	30	81.1	305	8	ADN99683	Adn99683	Novel hum	492	30	81.1	606	5	ABP43893	Arabidops
420	30	81.1	312	8	ADG63087	Adg63087	Transcrip	493	30	81.1	609	5	AAU78364	Arabidops
421	30	81.1	312	8	ADN72439	Adn72439	Thale cre	494	30	81.1	609	9	ADY30827	Thale cre
422	30	81.1	312	9	AEA26203	Aea26203	Stress to	495	30	81.1	612	6	AAO27114	qSH-1 gen
423	30	81.1	317	4	ABG03344	Abg03344	Novel hum	496	30	81.1	612	6	AAO27115	qSH-1 gen
424	30	81.1	318	4	ABG92714	Abg92714	Human pro	497	30	81.1	619	3	AAAG41273	Arabidops
425	30	81.1	322	6	ADA13325	Ada13325	Human int	498	30	81.1	623	4	AAE06678	Human nuc
426	30	81.1	330	6	ADA24352	Ada24352	Bacterial	499	30	81.1	623	7	ADC13533	Human NOV
427	30	81.1	344	6	ADA06235	Ada06235	Wheat cho	500	30	81.1	623	7	ADC31749	Human nov
428	30	81.1	344	8	ADOL7006	Adol7006	Wheat cho	501	30	81.1	625	8	ADS30749	Bacterial
429	30	81.1	346	3	AAAG21361	AAg21361	Arabidops	502	30	81.1	629	6	ABU34518	Protein e
430	30	81.1	346	8	ADX72771	Adx72771	Plant ful	503	30	81.1	634	8	ADQ18696	Human sof
431	30	81.1	353	7	ABM85588	Abm85588	Mouse pro	504	30	81.1	634	8	ADY22469	Plant ful
432	30	81.1	353	8	ADY08875	Ady08875	Plant ful	505	30	81.1	635	6	AAO26519	Protein o
433	30	81.1	354	8	ADN47354	Adn47354	Thermococ	506	30	81.1	638	8	ADX98261	Plant ful
434	30	81.1	364	3	AAAG14403	AAg14403	Arabidops	507	30	81.1	640	8	ADS30049	Bacterial
435	30	81.1	365	3	AAAG42876	AAg42876	Arabidops	508	30	81.1	641	3	AAAG41272	Arabidops
436	30	81.1	366	8	ADY06751	Ady06751	Plant ful	509	30	81.1	641	8	ADX92746	Plant ful
437	30	81.1	367	3	AAAG14402	AAg14402	Arabidops	510	30	81.1	643	6	ABU36780	Protein e
438	30	81.1	367	6	ADA83987	Ada83987	Human CDK	511	30	81.1	650	8	ADI82498	Human mod
439	30	81.1	368	3	AAAG42875	AAg42875	Arabidops	512	30	81.1	650	8	ADY12848	Plant ful
440	30	81.1	375	7	ADSO9101	Adso9101	Novel pro	513	30	81.1	674	4	AAW95253	Human pro
441	30	81.1	375	9	ADU40616	Adu40616	Novel hum	514	30	81.1	676	7	ADC33302	Human nov
442	30	81.1	384	8	ADU02313	Adu02313	Novel hum	515	30	81.1	703	8	ADH13194	Human mal
443	30	81.1	386	9	ADZ47759	Adz47759	Mycobacte	516	30	81.1	703	9	ADX07198	Cyclin-de
444	30	81.1	402	8	ADX91386	Adx91386	Plant ful	517	30	81.1	703	9	AEA15081	Human pol
445	30	81.1	405	3	AAAG21360	AAg21360	Arabidops	518	30	81.1	705	8	ADS28997	Bacterial
446	30	81.1	405	6	AAE36334	Aae36334	Arabidops	519	30	81.1	708	2	AAW36065	Human neu
447	30	81.1	414	8	AAAG21359	AAg21359	Arabidops	520	30	81.1	708	7	AAAG41271	Arabidops
448	30	81.1	414	8	ADJ48846	Adj48846	Oil-assoc	521	30	81.1	708	7	ADE56151	Human Pro
449	30	81.1	415	4	ABG30150	Abg30150	Novel hum	522	30	81.1	708	7	ADE56159	Human Pro
450	30	81.1	422	4	AAE01983	Aae01983	Human ATP	523	30	81.1	708	7	ADE56155	Human Pro
451	30	81.1	423	6	ABP99222	Abp99222	Orthosomy	524	30	81.1	708	8	ADJ64313	Cartilage
452	30	81.1	427	8	ADR86077	Adr86077	Aspergill	525	30	81.1	708	9	ADY14410	PRO polyp
453	30	81.1	428	4	ABG11496	Abg11496	Novel hum	526	30	81.1	712	8	ADX92580	Plant ful
454	30	81.1	440	4	ABG61159	Abg61159	Drosophil	527	30	81.1	718	4	AAU41732	Propionib
455	30	81.1	442	4	ABM69463	Abm69463	Drosophil	528	30	81.1	718	4	ABG03655	Novel hum
456	30	81.1	445	5	ABP69609	Abp69609	Human pol	529	30	81.1	718	6	ABM38251	Propionib
457	30	81.1	447	5	ABM05639	Abm05639	Hepatitis	530	30	81.1	735	4	ABM66233	Drosophil
458	30	81.1	449	2	AAW37129	Aaw37129	Hepatitis	531	30	81.1	739	6	ABU25591	Protein e
459	30	81.1	458	8	ADY08647	Ady08647	Plant ful	532	30	81.1	743	4	ABM59515	Drosophil
460	30	81.1	476	3	AAW08405	Aaw08405	Amino aci	533	30	81.1	753	5	AAU82710	Amino aci
461	30	81.1	476	4	ABM62927	Abm62927	Drosophil	534	30	81.1	761	5	ABP74116	Human TRI
462	30	81.1	480	6	ABG74402	Abg74402	Green pep	535	30	81.1	761	7	ADF74131	Human nov

536	30	81.1	763	2	AAW31852	Aaw31852 Mycobacte	609	30	81.1	1709	7	AD806701	Ade06701 Hepatitis
537	30	81.1	763	6	ABU48371	Abu48371 Protein e	610	30	81.1	1788	4	AU000016	Aau00016 Human Ple
538	30	81.1	779	8	ADT56006	Adt56006 Plant pol	611	30	81.1	1965	6	ABJ19786	Abj19786 DOCK 3 tu
539	30	81.1	810	9	ADZ22829	Adz22829 Muscle ca	612	30	81.1	1966	6	ABJ19797	Abj19797 DOCK 3 tu
540	30	81.1	810	3	AAE24089	Aae24089 Human pro	613	30	81.1	1966	6	ABJ19789	Abj19789 DOCK 3 tu
541	30	81.1	814	4	ABG30224	Abg30224 Novel hum	614	30	81.1	1966	6	ABJ19794	Abj19794 DOCK 3 tu
542	30	81.1	814	5	AAE26669	Aae26669 Human cad	615	30	81.1	1966	6	ABJ19795	Abj19795 DOCK 3 tu
543	30	81.1	814	7	ADD14141	Add14141 Human src	616	30	81.1	1966	6	ABJ19796	Abj19796 DOCK 3 tu
544	30	81.1	814	8	ADP26909	Adp26909 Human M-c	617	30	81.1	1966	6	ABJ19785	Abj19785 DOCK 3 tu
545	30	81.1	814	8	ABM81804	Abm81804 Tumour-as	618	30	81.1	1966	6	ABJ19787	Abj19787 DOCK 3 tu
546	30	81.1	817	6	ABR53281	AbR53281 Protein s	619	30	81.1	1966	6	ABJ19798	Abj19798 DOCK 3 tu
547	30	81.1	817	7	ADK63588	Adk63588 Disease t	620	30	81.1	1966	6	ABJ19791	Abj19791 DOCK 3 tu
548	30	81.1	817	8	ADN19160	Adn19160 Bacterial	621	30	81.1	1966	6	ABJ19792	Abj19792 DOCK 3 tu
549	30	81.1	819	6	ABJ19348	Abj19348 NOXV rela	622	30	81.1	1966	6	ABJ19793	Abj19793 DOCK 3 tu
550	30	81.1	819	8	ADO41696	Ado41696 Novel hum	623	30	81.1	1966	6	ABJ19788	Abj19788 DOCK 3 tu
551	30	81.1	823	4	ABO68711	AbO68711 Pseudomon	624	30	81.1	1966	6	ABJ19790	Abj19790 DOCK 3 tu
552	30	81.1	875	7	ABB71072	Abb71072 Drosophil	625	30	81.1	1981	8	ADSI11120	Adsi11120 Human the
553	30	81.1	875	8	ADJ40972	Adj40972 Fugu pher	626	30	81.1	1985	6	ABU09574	Abu09574 HCV Met-N
554	30	81.1	875	8	ADJ41023	Adj41023 Fugu pher	627	30	81.1	1985	6	ABU09575	Abu09575 HCV Met-N
555	30	81.1	881	8	ADS30366	Ads30366 Bacterial	628	30	81.1	1985	8	ADR38450	Adr38450 Hepatitis
556	30	81.1	942	8	ADQ66974	Adq66974 Novel hum	629	30	81.1	2065	4	ABE63705	AbE63705 Drosophil
557	30	81.1	953	9	ABM95482	Abm95482 M. xanthu	630	30	81.1	2091	8	ADN61449	Adn61449 Human KPP
558	30	81.1	972	6	ABU25429	Abu25429 Protein e	631	30	81.1	2132	9	ADV97875	Adv97875 Murine pr
559	30	81.1	995	4	ABBE2510	Abbe2510 Drosophil	632	30	81.1	2193	6	ABR42219	AbR42219 Human pro
560	30	81.1	1052	4	ABG27366	Abg27366 Novel hum	633	30	81.1	2201	2	AAW01680	Aaw01680 HCV NS2-N
561	30	81.1	1056	4	AAE01980	Aae01980 Human ATP	634	30	81.1	2217	9	ADX98203	Adx98203 Lysine de
562	30	81.1	1082	7	ADEL4368	Adel4368 Human int	635	30	81.1	2219	8	ADN00364	Adn00364 Novel hum
563	30	81.1	1085	8	ADS43269	Ads43269 Bacterial	636	30	81.1	2245	8	ADJ96651	Adj96651 Human NIM
564	30	81.1	1088	8	ADH61284	Adh61284 INTSIG pr	637	30	81.1	2382	6	ABP71619	Abp71619 Human WNK
565	30	81.1	1112	8	ADR66113	Adr66113 Human pro	638	30	81.1	2382	7	ADJ69598	Adj69598 Human hea
566	30	81.1	1112	8	ADR66455	Adr66455 Human pro	639	30	81.1	2382	9	ADX07450	Adx07450 Cyclin-de
567	30	81.1	1112	9	ADY18523	Ady18523 PRO poly	640	30	81.1	2382	9	ADX98202	Adx98202 Lysine de
568	30	81.1	1116	6	ABJ26002	Abj26002 Aspergill	641	30	81.1	2388	8	ADSI11119	Adsi11119 Human the
569	30	81.1	1116	6	ABU25402	Abu25402 Aspergill	642	30	81.1	3010	2	AAR20111	Aar20111 Non-A, no
570	30	81.1	1146	3	AAV92225	Aav92225 Human pat	643	30	81.1	3010	2	AAR20091	Aar20091 Non-A, no
571	30	81.1	1203	3	AAV43261	Aav43261 Human pat	644	30	81.1	3010	2	AAR34580	Aar34580 Human hep
572	30	81.1	1203	2	AAV28444	Aav28444 Human plc	645	30	81.1	3010	9	ADY06423	Ady06423 Non-A, no
573	30	81.1	1203	3	AAV92703	Aav92703 Human pat	646	30	81.1	3010	9	ADX40788	Adx40788 HCV polym
574	30	81.1	1203	5	AAE19829	Aae19829 Human pat	647	30	81.1	3010	9	ADX40806	Adx40806 HCV polym
575	30	81.1	1203	6	ABG74104	Abg74104 Human col	648	30	81.1	3011	2	AAR34468	Aar34468 Encoded b
576	30	81.1	1203	8	ADT07558	Adt07558 Human col	649	29	78.4	7	2	AAR84747	Aar84747 GST-dynam
577	30	81.1	1205	8	ADT07560	Adt07560 Human col	650	29	78.4	7	3	AAV79782	Aav79782 Proline-r
578	30	81.1	1213	8	ADT07556	Adt07556 Human col	651	29	78.4	7	3	AAV69980	Aav69980 Src SH3 r
579	30	81.1	1234	4	ABBS8721	Abbs8721 Drosophil	652	29	78.4	8	2	AAR84745	Aar84745 DYN domai
580	30	81.1	1234	4	ABBS8714	Abbs8714 Drosophil	653	29	78.4	9	2	AAR84746	Aar84746 Dynamlin r
581	30	81.1	1237	2	AAR13791	Aar13791 E75A prot	654	29	78.4	9	2	AAW99358	Aaw99358 Glycosyla
582	30	81.1	1237	6	AAE30117	Aae30117 Fruit fly	655	29	78.4	9	8	ADK88377	Adk88377 Human 191
583	30	81.1	1270	4	AAE01192	Aae01192 Human ATP	656	29	78.4	9	8	ADK88004	Adk88004 Human 191
584	30	81.1	1278	6	ABG74682	Abg74682 Human CGD	657	29	78.4	9	8	ADK88214	Adk88214 Human 191
585	30	81.1	1278	8	ADI82558	Adi82558 Human mod	658	29	78.4	9	8	ADK87474	Adk87474 Human 191
586	30	81.1	1345	5	AAE25097	Aae25097 Human kin	659	29	78.4	9	8	ADK87252	Adk87252 Human 191
587	30	81.1	1360	4	ABG30225	Abg30225 Novel hum	660	29	78.4	9	8	ADK83836	Adk83836 Human 191
588	30	81.1	1363	4	ABBE4266	Abbe4266 Drosophil	661	29	78.4	9	8	ADK87598	Adk87598 Human 191
589	30	81.1	1394	9	AEA32849	Aea32849 Modified	662	29	78.4	9	8	ADK88201	Adk88201 Human 191
590	30	81.1	1401	8	ADL13306	Adl13306 Human ste	663	29	78.4	9	8	ADK83353	Adk83353 Human 191
591	30	81.1	1401	9	ADX07508	Adx07508 Cyclin-de	664	29	78.4	9	8	ADK87960	Adk87960 Human 191
592	30	81.1	1419	8	ADN03630	Adn03630 Antipsori	665	29	78.4	9	8	ADK85972	Adk85972 Human 191
593	30	81.1	1426	4	AAE01984	Aae01984 Human ATP	666	29	78.4	9	8	ADK84372	Adk84372 Human 191
594	30	81.1	1426	5	ABP52159	Abp52159 Human 671	667	29	78.4	9	8	ADK86542	Adk86542 Human 191
595	30	81.1	1426	6	ABR54236	AbR54236 Human NOV	668	29	78.4	9	8	ADK84941	Adk84941 Human 191
596	30	81.1	1426	7	ADD37490	Add37490 Human tra	669	29	78.4	9	8	ADK85432	Adk85432 Human 191
597	30	81.1	1426	8	ADT27969	Adt27969 Human 671	670	29	78.4	9	8	ADK87823	Adk87823 Human 191
598	30	81.1	1426	9	ADY16951	Ady16951 PRO poly	671	29	78.4	9	8	ADK87316	Adk87316 Human 191
599	30	81.1	1426	9	ADY20464	Ady20464 PRO poly	672	29	78.4	9	8	ADK87734	Adk87734 Human 191
600	30	81.1	1426	9	ADY79907	Ady79907 Amino aci	673	29	78.4	9	8	ADK85509	Adk85509 Human 191
601	30	81.1	1445	8	ADQ66588	Adq66588 Novel hum	674	29	78.4	9	9	ADK87084	Adk87084 Human 191
602	30	81.1	1447	7	ABR57181	AbR57181 Drosophil	675	29	78.4	9	9	ADZ57264	Adz57264 Cytotoxic
603	30	81.1	1480	7	ADK63568	Adk63568 Disease-t	676	29	78.4	10	2	AAR93542	Aar93542 Random 10
604	30	81.1	1480	8	ADS43687	Ads43687 Bacterial	677	29	78.4	10	2	AAW70181	Aaw70181 N-termina
605	30	81.1	1538	4	AAE66466	Aae66466 Protein e	678	29	78.4	10	4	AAR05637	Aar05637 N-termina
606	30	81.1	1627	9	ABM92081	Abm92081 M. xanthu	679	29	78.4	10	6	ABO43501	AbO43501 M. tuberc
607	30	81.1	1647	9	ABBS8428	Abbs8428 Drosophil	680	29	78.4	10	7	ADF12073	Adf12073 Mycobacte
608	30	81.1	1709	7	ADE06699	Ade06699 Hepatitis	681	29	78.4	10	7	ADM40793	Adm40793 Mycobacte

682	29	78.4	10	8	ADK84706	Adk84706 Human 191	755	29	78.4	23	4	AAM29839	Peptide #
683	29	78.4	10	8	ADK85685	Adk85685 Human 191	756	29	78.4	23	4	ABB31143	Peptide #
684	29	78.4	10	8	ADK85777	Adk85777 Human 191	757	29	78.4	23	4	AAM69500	Human bon
685	29	78.4	10	8	ADK86765	Adk86765 Human 191	758	29	78.4	23	4	AAM57109	Human bra
686	29	78.4	10	8	ADK83626	Adk83626 Human 191	759	29	78.4	28	2	AAR84757	Dynamlin r
687	29	78.4	10	8	ADK89039	Adk89039 Human 191	760	29	78.4	29	2	AAR92880	Mycobacte
688	29	78.4	10	8	ADK85168	Adk85168 Human 191	761	29	78.4	29	2	AAR75582	M. tuberc
689	29	78.4	10	8	ADK89606	Adk89606 Human 191	762	29	78.4	29	6	AAW75582	M. tuberc
690	29	78.4	10	8	ADK89050	Adk89050 Human 191	763	29	78.4	30	2	ABG74421	M. tuberc
691	29	78.4	10	8	ADK89421	Adk89421 Human 191	764	29	78.4	30	2	AAW85676	45 kd M. t
692	29	78.4	10	8	ADK88753	Adk88753 Human 191	765	29	78.4	30	2	AAW18194	N-termina
693	29	78.4	10	8	ADK89234	Adk89234 Human 191	766	29	78.4	30	2	ABU56346	M. tuberc
694	29	78.4	10	8	ADK86210	Adk86210 Human 191	767	29	78.4	30	7	AAR39289	M. tuberc
695	29	78.4	10	8	ADK84193	Adk84193 Human 191	768	29	78.4	30	7	ADP45131	M. tuberc
696	29	78.4	10	8	ADK84663	Adk84663 Human 191	769	29	78.4	30	8	ADO36816	45kd majo
697	29	78.4	10	8	ADK84110	Adk84110 Human 191	770	29	78.4	30	8	ADR30590	Human IGF
698	29	78.4	10	8	ADK89040	Adk89040 Human 191	771	29	78.4	30	8	ADU64191	Mycobacte
699	29	78.4	11	2	AAW25461	AAW25461 SH3 domai	772	29	78.4	34	4	AAM30735	Peptide #
700	29	78.4	11	2	AAW25477	AAW25477 SH3 domai	773	29	78.4	34	4	ABB32023	Peptide #
701	29	78.4	12	2	AAW25470	AAW25470 SH3 domai	774	29	78.4	34	4	ABG70405	Human bon
702	29	78.4	12	2	AAW70109	AAW70109 Peptide p	775	29	78.4	34	4	AAM57975	Human bra
703	29	78.4	12	4	AAE05589	AAE05589 N-termina	776	29	78.4	34	4	AAM05854	Peptide #
704	29	78.4	12	4	AAE05570	AAE05570 N-termina	777	29	78.4	37	2	AAW12910	Fragile X
705	29	78.4	12	6	ABO43428	ABO43428 M. tuberc	778	29	78.4	39	5	ABW77846	Amino aci
706	29	78.4	12	6	ABO43452	ABO43452 M. tuberc	779	29	78.4	43	4	AAM20313	Peptide #
707	29	78.4	13	2	AAW84648	AAW84648 Grb2-SOS	780	29	78.4	43	4	ABB40781	Peptide #
708	29	78.4	14	2	AAW76702	AAW76702 N-termina	781	29	78.4	43	4	AM34547	Peptide #
709	29	78.4	14	2	AAW16953	AAW16953 Src SH3 d	782	29	78.4	43	4	ABW24982	Protein #
710	29	78.4	15	2	AAW38946	AAW38946 Peptide r	783	29	78.4	43	4	AAM74433	Human bon
711	29	78.4	15	8	ADK90145	Adk90145 Human 191	784	29	78.4	43	4	AAM61641	Human bra
712	29	78.4	15	8	ADK89791	Adk89791 Human 191	785	29	78.4	43	4	ABG56227	Human liv
713	29	78.4	15	8	ADK90444	Adk90444 Human 191	786	29	78.4	43	5	ABG44325	Human pep
714	29	78.4	15	8	ADK89709	Adk89709 Human 191	787	29	78.4	44	5	ABP28990	Streptoco
715	29	78.4	15	8	ADK90233	Adk90233 Human 191	788	29	78.4	47	7	ADH34263	Sterlet C
716	29	78.4	15	8	ADK90020	Adk90020 Human 191	789	29	78.4	48	7	ABM89624	Rice abio
717	29	78.4	15	8	ADK89961	Adk89961 Human 191	790	29	78.4	50	5	ABP34249	Human ORF
718	29	78.4	15	8	ADK89716	Adk89716 Human 191	791	29	78.4	52	4	AAU41114	Propionib
719	29	78.4	15	8	ADK90490	Adk90490 Human 191	792	29	78.4	52	6	ABM37633	Propionib
720	29	78.4	15	8	ADK90019	Adk90019 Human 191	793	29	78.4	54	4	AAU42763	Propionib
721	29	78.4	15	8	ADK90232	Adk90232 Human 191	794	29	78.4	54	6	ABM39282	Propionib
722	29	78.4	15	8	ADK90234	Adk90234 Human 191	795	29	78.4	58	5	ABP06010	Human ORF
723	29	78.4	15	8	ADK90388	Adk90388 Human 191	796	29	78.4	58	9	ADY64642	S. mangon
724	29	78.4	15	8	ADK90470	Adk90470 Human 191	797	29	78.4	60	4	AAW82646	Human imm
725	29	78.4	15	8	ADK89770	Adk89770 Human 191	798	29	78.4	60	4	ABG20790	Human hum
726	29	78.4	15	8	ADK90411	Adk90411 Human 191	799	29	78.4	64	8	ABO57564	Human gen
727	29	78.4	16	2	AAW25364	AAW25364 Ab1 SH3 d	800	29	78.4	67	8	ADK48110	Streptoco
728	29	78.4	16	6	AAE36353	AAE36353 Human BSD	801	29	78.4	70	6	ABM51031	Propionib
729	29	78.4	17	2	AAW05472	AAW05472 SH3-bindin	802	29	78.4	70	6	AAW75162	Propionib
730	29	78.4	17	2	AAW38051	AAW38051 PPPY mot	803	29	78.4	71	2	AAW75163	Human sec
731	29	78.4	17	2	AAW32331	AAW32331 Mycobacte	804	29	78.4	71	3	AAW76073	Human aki
732	29	78.4	17	2	AAW32400	AAW32400 Mycobacte	805	29	78.4	71	4	AAW56012	Human col
733	29	78.4	17	2	AAW64347	AAW64347 Mycobacte	806	29	78.4	71	4	AAW77018	Human coli
734	29	78.4	17	2	AAW81714	AAW81714 M. tuberc	807	29	78.4	71	5	ABW72212	Human pro
735	29	78.4	17	2	AAW81690	AAW81690 M. tuberc	808	29	78.4	71	6	ABO02039	Novel hum
736	29	78.4	17	2	AAW39001	AAW39001 M. tuberc	809	29	78.4	71	6	ABO02038	Novel hum
737	29	78.4	17	2	AAW38952	AAW38952 M. tuberc	810	29	78.4	71	9	ADZ12440	Human sec
738	29	78.4	17	2	AAW39144	AAW39144 M. tuberc	811	29	78.4	71	9	ADZ12439	Human sec
739	29	78.4	17	2	AAW39090	AAW39090 M. tuberc	812	29	78.4	73	4	AAU20656	Human sec
740	29	78.4	17	7	ADW49306	ADW49306 Novel WW	813	29	78.4	73	4	AAW91760	Human imm
741	29	78.4	19	2	AAW48258	AAW48258 Nucleolar	814	29	78.4	75	4	ABG26289	Novel hum
742	29	78.4	19	2	AAW96356	AAW96356 Nucleolar	815	29	78.4	76	9	ABW92044	M. xanthu
743	29	78.4	19	2	ABG92991	ABG92991 Nucleolar	816	29	78.4	77	3	AAW56033	Polypoli
744	29	78.4	19	5	ABP56586	ABP56586 Nucleus s	817	29	78.4	77	3	AAW89933	Human imm
745	29	78.4	19	8	ADU07176	ADU07176 Nucleolar	818	29	78.4	77	5	ABP34432	Human ORF
746	29	78.4	19	8	ADT61875	ADT61875 Human nuc	819	29	78.4	86	5	ABP32728	Human ORF
747	29	78.4	19	9	ADW81324	ADW81324 Intrabody	820	29	78.4	87	4	ABG28592	Novel hum
748	29	78.4	19	9	ADW88617	ADW88617 Nucleolus	821	29	78.4	87	8	ADG22729	Cyanophag
749	29	78.4	19	9	ADY32253	ADY32253 Novel can	822	29	78.4	90	3	ABW56590	Human pro
750	29	78.4	19	9	ABE17233	ABE17233 Human nuc	823	29	78.4	91	4	AAU21866	Human car
751	29	78.4	19	9	AEA43017	AEA43017 Nucleolar	824	29	78.4	91	7	ABE45834	Human car
752	29	78.4	20	2	AAW16984	AAW16984 ABL SH3 d	825	29	78.4	91	8	ADJ07252	Human car
753	29	78.4	20	2	AAW50220	AAW50220 ITPF1 frag	826	29	78.4	92	8	ABO57213	Human gen
754	29	78.4	22	9	AAW50220	AAW50220 ITPF1 frag	827	29	78.4	95	4	AAM23515	Human EST

828	29	78.4	95	4	AAM85416	Aam85416 Human imm	901	29	78.4	151	7	ADD38035	Add38035 Human sec
829	29	78.4	95	4	AAM95844	Aam95844 Human tes	902	29	78.4	153	2	AAM88293	Aam88293 Wheat gam
830	29	78.4	95	4	AB596375	Ab596375 Human tes	903	29	78.4	155	8	ADR10386	Adr10386 Human pro
831	29	78.4	98	4	AAO10464	Aao10464 Human pol	904	29	78.4	157	4	AAU62522	Aau62522 Propionib
832	29	78.4	105	4	ABB15328	Abb15328 Human ner	905	29	78.4	157	6	ABM59041	Abm59041 Propionib
833	29	78.4	105	9	AEA43772	Aea43772 Perhydrol	906	29	78.4	159	6	ABU33934	Abu33934 Protein e
834	29	78.4	107	9	ABM90928	Abm90928 M. xanthu	907	29	78.4	159	6	ABU34553	Abu34553 Protein e
835	29	78.4	110	8	ADM87746	Adm87746 Human EST	908	29	78.4	159	7	ADL36004	Adl36004 Human nov
836	29	78.4	112	4	AAU39507	Aau39507 Propionib	909	29	78.4	161	6	ABR56784	AbR56784 Human sec
837	29	78.4	112	6	ABM36026	Abm36026 Propionib	910	29	78.4	161	8	ADY05803	Ady05803 Plant ful
838	29	78.4	116	4	AAW19466	Aaw19466 Peptide #	911	29	78.4	162	3	AAI12442	Aai12442 Zea mays
839	29	78.4	116	4	ABB38933	Abb38933 Peptide #	912	29	78.4	162	4	AAU07336	Aau07336 1-aminocy
840	29	78.4	116	4	AAW32416	Aaw32416 Peptide #	913	29	78.4	162	3	AAU07336	Aau07336 1-aminocy
841	29	78.4	116	4	ABR23920	AbR23920 Protein #	914	29	78.4	162	6	ABU36794	Abu36794 Protein e
842	29	78.4	116	4	AAW72154	Aaw72154 Human bon	915	29	78.4	163	3	AAV65930	Aav65930 MUC1 muta
843	29	78.4	116	4	AAU46429	Aau46429 Propionib	916	29	78.4	166	6	ADA11893	Ada11893 Human nov
844	29	78.4	116	4	AAW59584	Aaw59584 Human bra	917	29	78.4	166	8	ADX67667	Adx67667 Plant ful
845	29	78.4	116	4	ABG53840	Abg53840 Human liv	918	29	78.4	167	2	AAV36228	Aav36228 Human sec
846	29	78.4	116	5	ABG41968	Abg41968 Human pep	919	29	78.4	167	5	ABP42995	Abp42995 Human ova
847	29	78.4	116	6	ABM42948	Abm42948 Propionib	920	29	78.4	167	6	ADA11722	Ada11722 Human nov
848	29	78.4	117	4	AAW84403	Aaw84403 Human imm	921	29	78.4	167	9	ABM97062	Abm97062 M. xanthu
849	29	78.4	117	4	AAO10964	Aao10964 Human pol	922	29	78.4	170	4	AAO13099	Aao13099 Human pol
850	29	78.4	119	5	ADK35157	Adk35157 Novel hum	923	29	78.4	172	3	AAV65932	Aav65932 MUC1 muta
851	29	78.4	121	4	AAV67557	Aav67557 Protein e	924	29	78.4	172	4	AAV52473	Aav52473 Mycobacte
852	29	78.4	121	4	ABG26404	Abg26404 Novel hum	925	29	78.4	172	9	ADM05076	Adm05076 C glutami
853	29	78.4	121	7	ADC333180	Adc333180 Human nov	926	29	78.4	172	9	ABE91615	Abe91615 Microbial
854	29	78.4	121	7	ADC333316	Adc333316 Human nov	927	29	78.4	173	3	AAV65933	Aav65933 MUC1 muta
855	29	78.4	123	3	ABG27165	Abg27165 Zea mays	928	29	78.4	173	8	ADX78231	Adx78231 Plant ful
856	29	78.4	124	3	ABM97215	Abm97215 M. xanthu	929	29	78.4	174	2	AAW64360	Aaw64360 Mycobacte
857	29	78.4	125	2	AAV10797	Aav10797 Amino aci	930	29	78.4	174	2	AAW81727	Aaw81727 M. tuberc
858	29	78.4	125	7	ABR47915	AbR47915 Novel hum	931	29	78.4	174	2	AAW39014	Aaw39014 M. tuberc
859	29	78.4	125	8	ADJ55470	Adj55470 Novel hum	932	29	78.4	174	2	AAV39157	Aav39157 M. tuberc
860	29	78.4	125	8	ADX87845	Adx87845 Plant ful	933	29	78.4	174	7	ABM88081	Abm88081 Rice abio
861	29	78.4	126	9	AEC04629	Aec04629 Human bre	934	29	78.4	175	8	ADN99298	Adn99298 Novel hum
862	29	78.4	127	4	AAO01196	Aao01196 Human pol	935	29	78.4	181	4	ABB50157	Abb50157 Human tra
863	29	78.4	127	4	AAU11220	Aau11220 Novel hum	936	29	78.4	183	3	AAQ08187	Aaq08187 Arabidops
864	29	78.4	130	5	ABR89427	AbR89427 Human pol	937	29	78.4	183	7	ABM73394	Abm73394 DNA clone
865	29	78.4	130	8	ADG65935	Adg65935 Novel hum	938	29	78.4	185	7	ABO71429	AbO71429 Pseudomon
866	29	78.4	132	3	AAV53421	Aav53421 Human col	939	29	78.4	185	7	ABO77768	AbO77768 Pseudomon
867	29	78.4	133	4	AAU50531	Aau50531 Propionib	940	29	78.4	186	8	ADX80177	Adx80177 Plant ful
868	29	78.4	133	4	ABG09963	Abg09963 Novel hum	941	29	78.4	187	2	AAW89878	Aaw89878 Antigen 1
869	29	78.4	133	5	ABU05968	Abu05968 M. tuberc	942	29	78.4	187	3	AAQ35655	Aaq35655 Arabidops
870	29	78.4	133	5	ABR98983	AbR98983 Endotheli	943	29	78.4	188	2	AAW95491	Aaw95491 M. tuberc
871	29	78.4	133	6	ABM47050	Abm47050 Propionib	944	29	78.4	189	4	AAU16302	Aau16302 Human nov
872	29	78.4	133	6	ABU37006	Abu37006 Protein e	945	29	78.4	189	4	AAU34636	Aau34636 E. coli c
873	29	78.4	133	6	ABU34926	Abu34926 Protein e	946	29	78.4	189	6	AAE330073	Aae330073 Arabidops
874	29	78.4	133	7	ABO68139	AbO68139 Pseudomon	947	29	78.4	189	6	ABU55371	Abu55371 Human nov
875	29	78.4	133	8	ADW05132	Adw05132 Human EDG	948	29	78.4	189	9	ABE78057	Abe78057 E. coli U
876	29	78.4	134	3	AGC26903	Agc26903 Zea mays	949	29	78.4	190	4	ABG08841	AbG08841 Novel hum
877	29	78.4	134	7	ABO68213	AbO68213 Pseudomon	950	29	78.4	190	7	ABO70602	AbO70602 Pseudomon
878	29	78.4	135	7	ADM06091	Adm06091 Human pro	951	29	78.4	191	7	ABO66030	AbO66030 Klebsiell
879	29	78.4	136	2	AAV12478	Aav12478 Human 5'	952	29	78.4	191	8	ADP04156	Adp04156 Human col
880	29	78.4	136	4	AAU51859	Aau51859 Propionib	953	29	78.4	193	5	ABJ10473	Abj10473 Breast ca
881	29	78.4	136	6	ABM48378	Abm48378 Propionib	954	29	78.4	193	5	ABP52841	Abp52841 Human ach
882	29	78.4	137	7	ABO81602	AbO81602 Pseudomon	955	29	78.4	193	5	AAU10337	Aau10337 Novel hum
883	29	78.4	137	4	AAU50475	Aau50475 Propionib	956	29	78.4	193	6	AAE33613	Aae33613 Human CAS
884	29	78.4	137	6	ABM46994	Abm46994 Propionib	957	29	78.4	193	7	ADJ68845	Adj68845 Human hea
885	29	78.4	138	4	AAO08702	Aao08702 Human pol	958	29	78.4	193	7	ADN39655	Adn39655 Cancer/an
886	29	78.4	138	6	ABM65442	Abm65442 Propionib	959	29	78.4	193	8	ADM69070	Adm69070 Human TAT
887	29	78.4	139	4	AAU64104	Aau64104 Propionib	960	29	78.4	193	9	ADX83185	Adx83185 Human TEG
888	29	78.4	139	6	ABM60623	Abm60623 Propionib	961	29	78.4	195	6	ABM70751	Abm70751 Photorhab
889	29	78.4	142	6	ADA33033	Ada33033 Acinetoba	962	29	78.4	200	8	ABO59883	AbO59883 Human gen
890	29	78.4	143	2	AAV36077	Aav36077 Extended	963	29	78.4	200	8	ADX87735	Adx87735 Plant ful
891	29	78.4	143	3	AAV65936	Aav65936 MUC1 muta	964	29	78.4	203	7	ABO82744	AbO82744 Pseudomon
892	29	78.4	143	3	AG44635	Ag44635 Zea mays	965	29	78.4	204	8	ADX97034	Adx97034 Plant ful
893	29	78.4	143	8	ADP19385	Adp19385 Human sec	966	29	78.4	206	8	ADX97250	Adx97250 Plant ful
894	29	78.4	144	7	ABO82491	AbO82491 Pseudomon	967	29	78.4	207	9	ABM91766	Abm91766 M. xanthu
895	29	78.4	146	8	ABO55205	AbO55205 Human gen	968	29	78.4	209	4	ABM67804	Abm67804 Drosophil
896	29	78.4	150	8	ADY12206	Ady12206 Plant ful	969	29	78.4	209	7	ABO63554	AbO63554 Klebsiell
897	29	78.4	151	6	ADA57406	Ada57406 Human sec	970	29	78.4	210	7	ABO63406	AbO63406 Klebsiell
898	29	78.4	151	6	ADA41281	Ada41281 Human sec	971	29	78.4	211	7	ADT88972	Adt88972 Tobacco M
899	29	78.4	151	6	ABR48027	AbR48027 Human sec	972	29	78.4	213	4	ABG14699	Abg14699 Novel hum
900	29	78.4	151	7	ADC74470	Adc74470 Human sec	973	29	78.4	213	7	ADB664730	AdB664730 Human pro

974 29 78.4 216 6 ADA48628
 975 29 78.4 216 7 ADJ11720
 976 29 78.4 216 7 ADJ11366
 977 29 78.4 216 7 AB075625
 978 29 78.4 217 5 ABJ10479
 979 29 78.4 217 6 AAE33619
 980 29 78.4 217 7 ABM88500
 981 29 78.4 219 3 AAG35854
 982 29 78.4 220 7 AB067993
 983 29 78.4 221 6 ADA48512
 984 29 78.4 221 7 ABR63182
 985 29 78.4 222 8 ADX91584
 986 29 78.4 228 7 ABM87943
 987 29 78.4 228 9 ABM96850
 988 29 78.4 229 8 ADT49803
 989 29 78.4 230 8 ADT59535
 990 29 78.4 231 4 AAB83411
 991 29 78.4 231 8 ADX71072
 992 29 78.4 232 4 ABG11714
 993 29 78.4 233 4 ABG16403
 994 29 78.4 233 6 ABUL19614
 995 29 78.4 235 6 ABR41859
 996 29 78.4 239 8 ADX88356
 997 29 78.4 242 8 ADX66752
 998 29 78.4 243 2 AAR58815
 999 29 78.4 243 3 AAG35853
 1000 29 78.4 243 8 ADN73199

ALIGNMENTS

RESULT 1
 AAB17240
 ID AAB17240 standard; peptide; 7 AA.

AC AAB17240;

DT 31-OCT-2000 (first entry)

DE SH3 antagonist peptide sequence SEQ ID NO:296.

KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
 KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
 KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
 KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;
 KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
 KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
 KW thrombosis; pharmaceutical.

XX Synthetic.

OS WO200024782-A2.

PN 04-MAY-2000.

PD 25-OCT-1999; 99WO-US025044.

PF 23-OCT-1998; 98US-0105371P.

PR 22-OCT-1999; 99US-00428082.

XX (AMGE-) AMGEN INC.

PI Feige U, Liu C, Cheetham J, Boone TC;

XX WPI; 2000-350702/30.

XX Novel composition of matter comprising an Fc domain and pharmacologically

PT active peptides, useful for treating cancer and autoimmune diseases.

XX Claim 39; Page 299; 608pp; English.

XX The present invention describes composition of matter (I) comprising an

CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
 CC (X1)a-P1-(X2)b, where: P1 = an Fc domain; X1 and X2 = are each
 CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
 CC (L2)d-P2-(L3)e-P3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
 CC P3, and P4 = are each independently sequences of pharmacologically active
 CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,
 CC c, d, e, and f = are each independently 0 or 1, provided that at least 1
 CC of a and b is 1. The composition can have cytostatic, antiasthmatic,
 CC thrombolytic and immunosuppressive activities. DNAs, vectors and host
 CC cells from the present invention can be used for producing pharmaceutical
 CC compositions. The compositions are useful for treating cancer, asthma,
 CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
 CC a Fab domain) can provide a longer half-life or incorporate functions
 CC such as Fc receptor binding, protein A binding, complement fixation, and
 CC possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to
 CC AAB18003 represent nucleotide and amino acid sequences used in the
 CC exemplification of the present invention

SQ Sequence 7 AA;

Query Match 100.0%; Score 37; DB 3; Length 7;
 Best Local Similarity 100.0%; Pred. No. 2e+06;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAPPVP 7

DB 1 ILAPPVP 7

RESULT 2

ABB73233

ID ABB73233 standard; peptide; 7 AA.

XX ABB73233;

DT 05-APR-2002 (first entry)

DE Src homology3 (SH3) antagonist peptide SEQ ID NO:296.

KW Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
 KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
 KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TNP;
 KW TPO mimetic peptide; EPO mimetic peptide; BMP; VEGF antagonist;
 KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
 KW cytostatic; antirheumatic; antiarthritis; antidiabetic; ophthalmological;
 KW antianaemic; anorectic; antiinfertility; haemostatic; dermatological;
 KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
 KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
 KW sleep disorder; neurological degenerative disease; anaemia;
 KW thrombocytopenia; metastatic tumour; systemic lupus erythematosus;
 KW Fanconi's syndrome.

XX Homo sapiens.

OS Synthetic.

PN WO200183525-A2.

PD 08-NOV-2001.

PF 02-MAY-2001; 2001WO-US014310.

XX 03-MAY-2000; 2000US-00563286.

XX (AMGE-) AMGEN INC.

PI Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;

XX WPI; 2002-130313/17.

XX Novel vehicle-peptide molecule or its multimers useful for treating
 PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
 PT diabetic retinopathy, obesity, sleep disorders and infertility.

PS Claim 39; Page 55; 176pp; English.

XX The present invention describes a vehicle-peptide molecule (I) or its
 CC cytostatic, antiinflammatory, antitumour, immunosuppressive,
 CC cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,
 CC antanaemic, anorectic, antifertility, haemostatic, dermatological and
 CC neuroprotective activities. (I) can be used as a therapeutic or
 CC prophylactic agent as well as for screening purposes. (I) is useful for
 CC diagnosing diseases characterised by dysfunction of their associated
 CC protein of interest, for identifying normal or abnormal proteins of
 CC interest, as a part of diagnostic kit to detect the presence of their
 CC proteins of interest in a biological sample. Additionally, (I) is useful
 CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,
 CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders, EPO-
 CC infertility, and neurological degenerative diseases. (I), comprising EPO-
 CC mimetic compounds are useful for treating disorders characterised by low
 CC red blood cell levels such as anaemia. The EPO-mimetic comprising
 CC compounds are useful for treating conditions that involve an existing
 CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet
 CC deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic
 CC tumour which result in thrombocytopaenia, systemic lupus erythematosus,
 CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABB35695 to ABB35777
 CC represent amino acid and nucleic acid sequences used in the
 CC exemplification of the present invention

XX Sequence 7 AA;

Query Match 100.0%; Score 37; DB 5; Length 7;
 Best Local Similarity 100.0%; Pred. No. 2e+06;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAPPVP 7
 DB 1 ILAPPVP 7

RESULT 3
 ADJ73387
 ID ADJ73387 standard; peptide; 7 AA.
 XX AC ADJ73387;
 XX DT 06-MAY-2004 (first entry)
 XX DE SH3 antagonist peptide sequence SeqID 842.
 KW mimetic; CDR mimetibody; gene therapy; neurogenic; immune;
 KW cardiovascular; infectious; malignant; neurologic disease; anaemia;
 KW immunomodulator; cardiant; antimicrobial; cytostatic; neuroprotective;
 KW SH3.
 XX OS Synthetic.
 XX PN WO2003084477-A2.
 XX PD 16-OCT-2003.
 XX PF 24-MAR-2003; 2003WO-US009139.
 XX PR 29-MAR-2002; 2002US-0368791P.
 XX PA (CENZ) CENTOCOR INC.
 XX PI Heavner GA, Knight DM, Scallion BJ, Ghayeb J;
 XX WPI; 2003-804237/75.
 XX DR New CDR mimetibody comprising a portion of a heavy or light chain
 PT variable region comprising human framework or ligand binding region,
 PT useful for preparing a composition for treating e.g., immune,
 PT cardiovascular or neurologic disease.
 XX PS Disclosure; SEQ ID NO 842; 97pp; English.

XX This invention relates to novel mammalian CDR mimetibodies, specific
 CC portions or variants thereof. Specifically, it refers to an antibody
 CC fragment where a protein has been inserted into, or replaces a portion
 CC of, one or more CDR regions, such that each CDR mimetibody comprises at
 CC least one portion of a heavy chain or light chain variable region, which
 CC itself comprises at least one human framework region and at least one
 CC ligand binding region (LBR). The present invention describes human
 CC mimetibodies, including modified immunoglobulins and cleavage products
 CC that can be useful in gene therapy and the generation of transgenic
 CC plants and animals. Furthermore, the CDR mimetibody is useful for
 CC preparing compositions for modulating, treating or reducing the symptoms
 CC of immune, cardiovascular, infectious, malignant and/or neurologic
 CC diseases, as well as anaemia. Accordingly, they exhibit immunomodulator,
 CC cardiant, antimicrobial, cytostatic and neuroprotective activities. This
 CC peptide sequence is an SH3 antagonist peptide sequence used to make a
 CC mimetibody of the invention.

XX Sequence 7 AA;

Query Match 100.0%; Score 37; DB 7; Length 7;
 Best Local Similarity 100.0%; Pred. No. 2e+06;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAPPVP 7
 DB 1 ILAPPVP 7

RESULT 4
 ADJ53021
 ID ADJ53021 standard; peptide; 7 AA.
 XX AC ADJ53021;
 XX DT 06-MAY-2004 (first entry)
 XX DE CH1 deleted mimetibody-related peptide SeqID842.
 KW CH1 deleted mimetibody; immunosuppressive; cardiovascular; cardiant;
 KW hypotensive; neuroprotective; nootropic; antibacterial; virucide;
 KW fungicide; gene therapy; immune disorder; cardiovascular disease;
 KW arrhythmia; hypertension; heart failure; neurodegenerative;
 KW multiple sclerosis; dementia; Alzheimer's disease; anaemia;
 KW cancerous condition; infectious disease; bacterial infection;
 KW viral infection; fungal infection.
 XX OS Unidentified.
 XX OS Synthetic.
 XX PN WO2004002417-A2.
 XX PD 08-JAN-2004.
 XX PF 27-JUN-2003; 2003WO-US020347.
 XX PR 28-JUN-2002; 2002US-0392431P.
 XX PA (CENZ) CENTOCOR INC.
 XX PI Heavner GA, Knight DM, Ghayeb J, Scallion BJ, Nesspor TC;
 XX PI Kutoloshki KA;
 XX WPI; 2004-082870/08.
 XX DR New CH1-deleted mimetibody polypeptides and nucleic acids, useful for
 PT modulating, treating, alleviating, preventing an immune, cardiovascular,
 PT or neurodegenerative disease or disorder, anemia, cancer, or infectious
 PT diseases.
 XX Claim 3; SEQ ID NO 842; 129pp; English.
 XX PS This invention relates to CH1 deleted mimetibodies (and the DNA sequences

CC which encode them), compositions, methods and uses. The invention may be
 CC useful for the development of compounds with an immunosuppressive,
 CC cardiovascular, cardiac, hypotensive, neuroprotective, neurotropic,
 CC antibacterial, virucide or fungicide activity. In addition, the disclosed
 CC sequences may prove useful for gene therapy. The CHI-deleted mimetibody
 CC is useful for diagnosing or treating a disease condition in a cell,
 CC tissue, organ or animal, specifically for modulating, treating,
 CC alleviating, preventing the incidence or reducing the symptoms of an
 CC immune, cardiovascular (for example arrhythmia, hypertension or heart
 CC failure), or neurodegenerative (for example multiple sclerosis, dementia
 CC or Alzheimer's disease) diseases or disorders, anaemia, cancerous
 CC conditions, or infectious diseases (for example bacterial, viral or
 CC fungal infection). The present sequence is that of a peptide which may be
 CC used during the creation of a mimetibody of the invention.

XX Sequence 7 AA;

Query Match 100.0%; Score 37; DB 8; Length 7;
 Best Local Similarity 100.0%; Pred. No. 2e+06;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 ILAPPVP 7
 Db 1 ILAPPVP 7

RESULT 5
 ADJ51982
 ID ADJ51982 standard; peptide; 7 AA.

AC ADJ51982;

DT 06-MAY-2004 (first entry)

DE CHI deleted mimetibody-related peptide SeqID842.

XX CHI deleted mimetibody; osteopathic; cardiovascular-Gen;
 KW dermatological-Gen; auditory; endocrine-Gen; gastrointestinal-Gen;
 KW gynaecological-Gen; hepatotropic; haemostatic; immunomodulator;
 KW antiallergic; muscular-Gen; cytostatic; antiinflammatory; neuroleptic;
 KW ophthalmologic; nephrotropic; respiratory-Gen; tumour necrosis factor;
 KW TNF; cytokine; bone disorder; joint disorder; cardiovascular disorder;
 KW dental disorder; oral disorder; dermatological disorder; ear disorder;
 KW nose disorder; throat disorder; endocrine disorder; metabolic disorder;
 KW gastrointestinal disorder; gynaecological disorder; hepatic disorder;
 KW obstetric disorder; haematologic disorder; immunological disorder;
 KW allergic disorder; infectious disorder; musculoskeletal disorder;
 KW oncological disorder; neurologic disorder; nutritional disorder;
 KW ophthalmologic disorder; pediatric disorder; psychiatric disorder;
 KW renal disorder; pulmonary disorder.

XX Unidentified.

OS Synthetic.

PN WO2004002424-A2.

XX 08-JAN-2004.

XX 30-JUN-2003; 2003WO-US020495.

XX 28-JUN-2002; 2002US-0392431P.

PR 19-SEP-2002; 2002US-0412144P.

XX (CENZ) CENTOCOR INC.

XX Heavner GA, Knight DM, Ghayab J, Scallion BJ, Nesspor TC;

PI Kutloski KA;

XX WPI; 2004-082872/08.

XX New CHI deleted mimetibody polypeptide and nucleic acid, useful for
 PT diagnosing, preventing or treating cardiovascular, dermatologic,
 PT endocrine, gastrointestinal, gynecologic, infectious, neurologic and

PT nutritional disorders.

XX Claim 15; SEQ ID NO 842; 123pp; English.

XX This invention relates to CHI deleted mimetibodies (and the DNA sequences
 CC which encode them), compositions, methods and uses. The invention may be
 CC useful for the development of compounds with an osteopathic,
 CC cardiovascular-Gen, dermatological-Gen, auditory, endocrine-Gen,
 CC gastrointestinal-Gen, gynaecological-Gen, hepatotropic, haemostatic,
 CC immunomodulator, antiallergic, ophthalmologic, nephrotropic or
 CC antiinflammatory, neuroleptic, musculoskeletal, neurologic or
 CC respiratory-Gen activity acting as a tumour necrosis factor (TNF)-
 CC modulator or cytokine-agonist. The methods and compositions of the
 CC present invention are useful for the diagnosis, prevention and/or
 CC treatment of diseases or conditions associated with aberrant expression
 CC or activity of the CHI deleted mimetibody, such as a bone or joint,
 CC cardiovascular, dental or oral, dermatological, ear, nose or throat,
 CC endocrine, metabolic, gastrointestinal, gynaecological, hepatic,
 CC obstetric, haematologic, immunological, allergic, infectious,
 CC musculoskeletal, oncological, neurological, nutritional, ophthalmologic,
 CC pediatric, psychiatric, renal or pulmonary disorders. The present
 CC sequence is that of a peptide which may be used during the creation of a
 CC mimetibody of the invention.

XX Sequence 7 AA;

Query Match 100.0%; Score 37; DB 8; Length 7;
 Best Local Similarity 100.0%; Pred. No. 2e+06;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 ILAPPVP 7
 Db 1 ILAPPVP 7

RESULT 6

AAW05414

ID AAW05414 standard; peptide; 13 AA.

XX AAW05414;

XX 24-FEB-1998 (first entry)

DE Src SH3 domain-binding peptide, T12SRC.1.

XX Src-homology region 3 domain; human; mouse; SH3 domain; cell growth;
 KW cellular signalling element; cellular structural element; malignancy;
 KW protein identification; functional domain; protein screening;
 KW cellular signal transduction process; binding peptide.

XX Homo sapiens.

OS Synthetic.

XX Key Location/Qualifiers

FT Modified-site 1 /note= "Biotin labelled"

FT Modified-site 13 /note= "C-terminal amide"

XX WO9631625-A1.

PD 10-OCT-1996.

XX 04-APR-1996; 96WO-US004454.

XX 07-APR-1995; 95US-00417872.

PR 03-APR-1996; 96US-00630915.

XX (CYTO-) CYTOGEN CORP.

PA (UYNC-) UNIV NORTH CAROLINA.

XX Sparks AB, Hoffman N, Kay BK, Fowlkes DM, McConnell SJ;

```

DR WPI; 1996-465045/46.
XX
XX Identifying polypeptide(s) having specific functional domain (esp. SH3
PT domain) - comprises detecting selective binding to recognition unit,
PT regardless of sequence homology.
XX
XX Example; Page 81; 174pp; English.
XX
XX AAW05414 and AAW05415 represent variants of the Src Src-homology region 3
CC (SH3) domain-binding peptide termed pSrcCII (see AAW05412). These
CC sequences were used to probe human cDNA libraries to identify human SH3
CC domain containing proteins (such as AAW05400), that can be used in the
CC method of the invention. The method of the invention is for identifying
CC polypeptides containing functional domains of interest (especially SH3
CC domains). It comprises contacting a multivalent recognition unit (RU)
CC complex with a number of peptides and identifying polypeptides having a
CC selective binding affinity for the RU complex. The method is based on
CC functional similarities and does not rely on sequence similarities. Prior
CC methods only gave limited success for identifying proteins containing an
CC SH3 domain due to the minimal sequence homology among known SH3 proteins.
CC Multivalent RU complexes are particularly suited to screening for
CC polypeptides containing functional domains that are similar to, but not
CC identical in sequence to, the original target functional domain. The new
CC method enables proteins having a common function to be identified.
CC Identification of novel SH3 proteins will be useful for a better
CC understanding of cell growth, malignancy, signal transduction processes,
CC etc. New candidate drugs can be identified, and their specificities (e.g.
CC pharmacological activities) can be assessed using the method of the
CC invention
XX
XX Sequence 13 AA;
SQ
Query Match 100.0%; Score 37; DB 2; Length 13;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAPPVP 7
Db |||||
2 ILAPPVP 8

RESULT 7
ID AAW05481
XX AAW05481 standard; peptide; 13 AA.
AC AAW05481;
XX
XX 24-FEB-1998 (first entry)
DT
XX
XX SH3-binding peptide T12SRC1.
DE
XX
XX Src-homology region 3 domain; human; mouse; SH3 domain; cell growth;
KW cellular signalling element; cellular structural element; malignancy;
KW protein identification; functional domain; protein screening;
KW cellular signal transduction process; binding peptide.
XX
XX Synthetic.
OS
XX
XX WO9631625-A1.
PN
XX
XX 10-OCT-1996.
PD
XX
XX 04-APR-1996; 96WO-US004454.
PF
XX
XX 07-APR-1995; 95US-00417872.
PR
XX
XX 03-APR-1996; 96US-00630915.
XX
XX (CYTO-) CYTOGEN CORP.
PA (UYNC-) UNIV NORTH CAROLINA.
XX
XX Sparks AB, Hoffman N, Kay BK, Fowlkes DM, McConnell SJ;
PI
XX WPI; 1996-465045/46.
DR

XX Identifying polypeptide(s) having specific functional domain (esp. SH3
PT domain) - comprises detecting selective binding to recognition unit,
PT regardless of sequence homology.
XX
XX Example; Fig 13; 174pp; English.
XX
XX AAW05445-W05492 represent Src-homology region 3 (SH2) domain binding
CC peptides. These sequences were used as parts of multivalent recognition
CC unit complexes used in the method of the invention. The method of the
CC invention is for identifying polypeptides containing functional domains
CC of interest (especially SH3 domains). It comprises contacting a
CC multivalent recognition unit (RU) complex with a number of peptides and
CC identifying polypeptides having a selective binding affinity for the RU
CC complex. The method is based on functional similarities and does not rely
CC on sequence similarities. Prior methods only gave limited success for
CC identifying proteins containing an SH3 domain due to the minimal sequence
CC homology among known SH3 proteins. Multivalent RU complexes are
CC particularly suited to screening for polypeptides containing functional
CC domains that are similar to, but not identical in sequence to, the
CC original target functional domain. The new method enables proteins having
CC a common function to be identified. Identification of novel SH3 proteins
CC will be useful for a better understanding of cell growth, malignancy,
CC signal transduction processes, etc. New candidate drugs can be
CC identified, and their specificities (e.g. pharmacological activities) can
CC be assessed using the method of the invention
XX
XX Sequence 13 AA;
SQ
Query Match 100.0%; Score 37; DB 2; Length 13;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAPPVP 7
Db |||||
2 ILAPPVP 8

RESULT 8
ID AAW1100
XX AAW1100 standard; peptide; 13 AA.
XX
XX AAW1100;
AC
XX
XX 25-JUN-1997 (first entry)
DT
XX
XX Src SH3 domain-binding peptide used in signal transduction modulation.
DE
XX
XX Src; SH3; Src homology region 3; binding affinity; oncogenic protein;
KW protein tyrosine kinase; signal transduction; RNA processing;
KW trafficking; translation.
XX
XX Synthetic.
OS
XX
XX WO9603649-A1.
PN
XX
XX 08-FEB-1996.
PD
XX
XX 24-JUL-1995; 95WO-US009382.
PF
XX
XX 22-JUL-1994; 94US-00278865.
PR
XX
XX 07-JUN-1995; 95US-00483555.
XX
XX (UYNC-) UNIV NORTH CAROLINA.
PA
XX
XX Sparks AB, Kay BK, Thorn JM, Quilliam LA, Der CJ;
PI
XX WPI; 1996-117151/12.
DR
XX
XX Peptide with binding affinity for Src homology region 3 (SH3) domains of
PT proteins - useful for e.g. modulating signal transduction pathways at the
PT cellular level, esp. protein tyrosine kinase-mediated.
PT
XX

```

PS Claim 34; Page 80; 116pp; English.

XX AAW1098-W11124 are peptides that bind to the Src SH3 domain. The SH3
CC binding peptides are useful in modulating signal transduction pathways at
CC the cellular level (especially protein tyrosine kinase-mediated),
CC modulating oncogenic protein activity, or providing compounds for the
CC development of drugs with the ability to modulate broad classes, as well
CC as specific classes, of proteins involved in signal transduction and also
CC for regulating the processing, trafficking or translation of RNA.
CC Conjugates of the peptides with detectable labels or imaging agents are
CC useful for imaging cells, tissues and organs in which Src or Src-related
CC proteins are expressed

XX Sequence 13 AA;

Query Match 100.0%; Score 37; DB 2; Length 13;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAPPVP 7
Db 2 ILAPPVP 8
|||||

RESULT 9
AAW1117
ID AAW1117 standard; peptide; 13 AA.
XX AC AAW1117;

DT 25-JUN-1997 (first entry)

XX Src SH3 domain-binding peptide used in signal transduction modulation.
KW Src; SH3; Src homology region 3; binding affinity; oncogenic protein;
KW protein tyrosine kinase; signal transduction; RNA processing;
KW trafficking; translation.

OS Synthetic.

XX WO9603649-A1.

XX 08-FEB-1996.

XX 24-JUL-1995; 95WO-US009382.

XX 22-JUL-1994; 94US-00278865.

XX 07-JUN-1995; 95US-00483555.

XX (UYNC-) UNIV NORTH CAROLINA.

XX Sparks AB, Kay BK, Thorn JM, Quilliam LA, Der CJ;

XX WPI; 1996-117151/12.

XX Peptide with binding affinity for Src homology region 3 (SH3) domains of
PT proteins - useful for e.g. modulating signal transduction pathways at the
PT cellular level, esp. protein tyrosine kinase-mediated.

PS Claim 39; Page 85; 116pp; English.

XX AAW1098-W11124 are peptides that bind to the Src SH3 domain. The SH3
CC binding peptides are useful in modulating signal transduction pathways at
CC the cellular level (especially protein tyrosine kinase-mediated),
CC modulating oncogenic protein activity, or providing compounds for the
CC development of drugs with the ability to modulate broad classes, as well
CC as specific classes, of proteins involved in signal transduction and also
CC for regulating the processing, trafficking or translation of RNA.
CC Conjugates of the peptides with detectable labels or imaging agents are
CC useful for imaging cells, tissues and organs in which Src or Src-related
CC proteins are expressed

XX Sequence 13 AA;

Query Match 100.0%; Score 37; DB 2; Length 13;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAPPVP 7
Db 4 ILAPPVP 10
|||||

RESULT 10
AAW38066
ID AAW38066 standard; peptide; 13 AA.
XX AC AAW38066;

DT 23-APR-1998 (first entry)

XX PPPY motif containing peptide used to bind WW domains.

XX Peptide recognition unit; WW domain; cell signalling; growth regulation;
KW cytoskeleton organisation; targeted drug screening; modulator;
KW WW domain interaction; YAP protein; dystrophin.

OS Synthetic.

XX WO9737223-A1.

XX 09-OCT-1997.

XX 03-APR-1997; 97WO-US005547.

XX 03-APR-1996; 96US-00630916.

XX (CYTO-) CYTOGEN CORP.

XX (UYNC-) UNIV NORTH CAROLINA.

XX Pirozzi G, Kay BK, Fowlkes DM;

XX WPI; 1997-503234/46.

XX Identifying cell signalling and growth regulatory polypeptides by
PT reaction with multivalent recognition complex - polypeptides are useful
PT in targeted drug selection.

PS Disclosure; Fig 15A; 220pp; English.

XX Peptides AAW38057-67 contain PPPY-like motifs. The PPPY motif is found
CC in the proline rich regions of Wap-1 and Wap-2 proteins. Peptides
CC containing this residue have been shown to bind the YAP WW domain, but
CC not the WW domain from dystrophin or to a panel of SH3 domains. Peptides
CC AAW38057-67 were biotinylated and complexed with alkaline streptavidin,
CC and used in a cross affinity mapping experiment. They were tested for
CC their ability to bind to the 12 individual novel WW domains of WWP1
CC (AAW36794), WWP2 (AAW37695), WWP3 (AAW37696) and WWP4 (AAW36797), which
CC were expressed as glutathione-S-transferase expression proteins. The
CC present peptide, derived from Src protein, does not bind the WW domains
CC of the novel proteins. The WW domain is a small functional domain. Its
CC name is derived from the observation that two tryptophan residues, one in
CC the amino terminal portion of the WW domain and one in the carboxyl
CC terminal portion, are conserved. Most proteins containing WW domains have
CC a function involving cell signalling and growth regulation or the
CC organisation of the cytoskeleton. Polypeptides containing a WW domain are
CC identified by treating a multivalent recognition unit complex that has
CC selective binding affinity for a WW domain, with many polypeptides and
CC identifying those with selective affinity for the complex. Proteins
CC containing WW domains are used for targeted drug screening, i.e. to
CC identify potential modulators of specific WW domain interactions

XX Sequence 13 AA;

Query Match 100.0%; Score 37; DB 2; Length 13;
Best Local Similarity 100.0%; Pred. No. 28;


```

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAPPVP 7
DB 2 ILAPPVP 8

RESULT 11
AAW25513
ID AAW25513 standard; peptide; 13 AA.
XX
AC AAW25513;
XX
DT 27-MAR-1998 (first entry)
XX
DE SH3 synthetic binding peptide.
XX
KW Cortactin; SH3 domain; binding peptide; Src homology region 3;
KW tyrosine kinase; immune response; lymphokine; interleukin 1; Nck; Abl;
KW PLCgamma; p53bp2; Crk; Yes; Grb2.
XX
OS Synthetic.
XX
PN WO9730074-A1.
XX
PD 21-AUG-1997.
XX
PF 14-FEB-1997; 97WO-US002298.
XX
PR 16-FEB-1996; 96US-00602999.
XX
PA (CYTO-) CYTOGEN CORP.
PA (UINC-) UNIV NORTH CAROLINA.
XX
PI Sparks AB, Kay BK, Thorn JM, Quilliam LA, Der CJ, Fowlkes DM;
PI Rider JB;
XX
DR WPI; 1997-424972/39.
XX
PT Src homology region 3 binding peptide - used to activate Src tyrosine
PT kinase(s) and to stimulate immune response by increasing production of
PT certain lymphokine(s), e.g. interleukin-1.
XX
PS Disclosure; Fig 7; 131pp; English.
XX
CC The present sequence represents a Src SH3 synthetic binding peptide. SH3
CC (Src homology region 3) binding peptides are selected from: (a) peptides
CC which bind the SH3 domain of Cortactin; (b) peptides which bind the
CC middle SH3 domain of Nck; (c) peptides which bind the SH3 domain of Abl;
CC (d) peptides which bind the SH3 domain of Src; (e) peptides which bind
CC the SH3 domain of PLC gamma; (f) peptides which bind the SH3 domain of
CC p53bp2; (g) peptides which bind the amino-terminal SH3 domain of Crk; (h)
CC peptides which bind the SH3 domain of Yes; and (i) peptides which bind
CC the amino-terminal SH3 domain of Grb2. The purified binding peptides can
CC be used in the method to identify inhibitors of their binding to their
CC respective SH3 domains, which could be used to modulate the
CC pharmacological activity of proteins or polypeptide containing the SH3
CC domain. The peptides can also be used to activate Src or Src-related
CC protein tyrosine kinases, to stimulate the immune response by increasing
CC the production of certain lymphokines, e.g. tumour necrosis factor-alpha
CC and interleukin-1, or to deliver a conjugated molecule to certain
CC cellular compartments containing Src or Src related proteins
XX
SQ Sequence 13 AA;
Query Match 100.0%; Score 37; DB 2; Length 13;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAPPVP 7
DB 2 ILAPPVP 8

RESULT 12
ADB49256
ID ADB49256 standard; peptide; 13 AA.
XX
AC ADB49256;
XX
DT 04-DEC-2003 (first entry)
XX
DE Biotinylated WW domain binding peptide #37.
XX
KW WW domain; drug candidate screening; drug discovery; drug modification;
KW drug refinement; immunogen; WW binding protein; WW domain.
XX
OS Unidentified.
XX
PN US2003077577-A1.
XX
PD 24-APR-2003.
XX
PF 28-JUN-2002; 2002US-00185050.
XX
PR 03-APR-1996; 96US-00630916.
PR 03-APR-1997; 97US-00826516.
XX
PA (PIRO/) PIROZZI G.
PA (KAYB/) KAY B K.
PA (FOWL/) FOWLKES D M.
XX
PI Pirozzi G, Kay BK, Fowlkes DM;
XX
DR WPI; 2003-635075/60.
XX
PT Novel purified polypeptide comprising WW domain, useful for drug
PT discovery, modification and refinement, for discovering polypeptides
PT involved in pharmacological activities, or as an immunogen to generate
PT antibodies.
XX
PS Example; Fig 15A; 133pp; English.
XX
CC The invention describes a purified polypeptide (I) comprising a WW domain
CC which has a sequence (S1) selected from 11 sequences fully defined in the
CC specification, a sequence (S2) selected from 48 sequences fully defined
CC in the specification or a sequence (S3) comprising 683, 906, 224 or 725
CC amino acids fully defined in the specification. (I) is useful for
CC screening a potential drug candidate, by allowing (I) to come into
CC contact with at least one recognition unit having a selective affinity
CC for the WW domain in (I), in the presence of an amount of a potential
CC drug candidate, such that (I) and the recognition unit are capable of
CC interacting when brought into contact with one another in the absence of
CC the drug candidate, and determining the effect, if any, of the presence
CC of the amount of the drug candidate on the interaction of (I) with the
CC recognition unit. (I) is useful for drug discovery, modification and
CC refinement, for discovering polypeptides involved in pharmacological
CC activities, or as an immunogen to generate antibodies. This is the amino
CC acid sequence of a WW domain binding peptide.
XX
SQ Sequence 13 AA;
Query Match 100.0%; Score 37; DB 7; Length 13;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAPPVP 7
DB 2 ILAPPVP 8

RESULT 13
AAAY24028
ID AAAY24028 standard; peptide; 16 AA.
XX
AC AAAY24028;

```

```

XX DT 29-SEP-1999 (first entry)
XX DE Synthetic peptide comprising a PDZ-binding domain.
XX KW Human; MMS1 protein; MMS1 interacting protein; tumour suppression;
XX KW MMS1 pathway; immunogen; cancer; cell neoplastic growth.
XX OS Synthetic.
XX PN WO9336566-A1.
XX PD 22-JUL-1999.
XX PF 19-JAN-1999; 99WO-US000995.
XX PR 20-JAN-1998; 98US-0071861P.
XX PA (MYRI-) MYRIAD GENETICS INC.
XX PI Bartel PL, Tavtigian SV;
XX DR WPI; 1999-458472/38.
XX PT MMS1, an MMS1 (tumour suppressor) interacting protein and related
XX PT polynucleotides.
XX PS Example 4; Page 48; 107pp; English.
XX CC The present sequence represents a PDZ-binding domain peptide, used in the
XX CC course of the invention. The specification describes a MMS1 protein, a
XX CC MMS1 interacting protein involved in tumour suppression activity in the
XX CC MMS1 pathway. MMS1, antigenic fragments or fusion proteins of these are
XX CC used as immunogens for antibody production. Primers derived from MMS1
XX CC genomic clones can be used for identification of MMS1 genes and for
XX CC synthesis by amplification of MMS1 DNA or RNA. Detecting an alteration
XX CC in MMS1 can be used to diagnose cancer. A germline alteration in an
XX CC MMS1 gene is indicative of a predisposition to cancer. A somatic
XX CC mutation in an MMS1 gene is indicative that the tissue is cancerous.
XX CC Analysis of MMS1 and MMS1 (or PDZ domain 6 of MMS1) binding
XX CC interactions can be used for detection of alterations in MMS1 associated
XX CC with cancer. Wild-type MMS1 or a homologue can be used to supply wild-
XX CC type MMS1 gene function (or a substantially similar function) to a cell,
XX CC which has lost the gene function due to a MMS1 gene mutation. The gene
XX CC suppresses neoplastic growth of the cell. Transgenic animals having an
XX CC altered MMS1 can be used as a model for identifying drug candidates
XX CC useful in treating cancer
XX SQ Sequence 16 AA;
XX Query Match 100.0%; Score 37; DB 2; Length 16;
XX Best Local Similarity 100.0%; Pred. No. 34;
XX Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX QY 1 ILAPPVP 7
XX Db |||||
XX 5 ILAPPVP 11
XX RESULT 14
XX AA05412
XX ID AA05412 standard; peptide; 17 AA.
XX AC AA05412;
XX XX
XX DT 24-FEB-1998 (first entry)
XX XX
XX DE Src SH3 domain-binding peptide, pSrcCII.
XX KW Src-homology region 3 domain; human; mouse; SH3 domain; cell growth;
XX KW cellular signalling element; cellular structural element; malignancy;
XX KW protein identification; functional domain; protein screening;
XX KW cellular signal transduction process; binding peptide.

```

```

XX OS Homo sapiens.
XX PH Key Location/Qualifiers
XX FT Modified-site 1
XX FT Modified-site 17 /note= "Biotin labelled"
XX FT Modified-site 17 /note= "C-terminal amide"
XX PN WO9631625-A1.
XX PD 10-OCT-1996.
XX PF 04-APR-1996; 96WO-US004454.
XX PR 07-APR-1995; 95US-00417872.
XX PR 03-APR-1996; 96US-00630915.
XX PA (CYTO-) CYTOGEN CORP.
XX PA (UYNC-) UNIV NORTH CAROLINA.
XX PI Sparks AB, Hoffman N, Kay BK, Fowlkes DM, McConnell SJ;
XX DR WPI; 1996-465045/46.
XX PT Identifying polypeptide(s) having specific functional domain (esp. SH3
XX PT domain) - comprises detecting selective binding to recognition unit,
XX PT regardless of sequence homology.
XX PS Example; Page 76; 174pp; English.
XX CC This sequence represents the Src Src-homology region 3 (SH3) domain-
XX CC binding peptide termed pSrcCII. pSrcCII was used to identify human and
XX CC mouse SH3 domain containing proteins (such as AAW05386), that can be used
XX CC in the method of the invention. The method of the invention is for
XX CC identifying polypeptides containing functional domains of interest
XX CC (especially SH3 domains). It comprises contacting a multivalent
XX CC recognition unit (RU) complex with a number of peptides and identifying
XX CC polypeptides having a selective binding affinity for the RU complex. The
XX CC method is based on functional similarities and does not rely on sequence
XX CC similarities. Prior methods only gave limited success for identifying
XX CC proteins containing an SH3 domain due to the minimal sequence homology
XX CC among known SH3 proteins. Multivalent RU complexes are particularly
XX CC suited to screening for polypeptides containing functional domains that
XX CC are similar to, but not identical in sequence to, the original target
XX CC functional domain. The new method enables proteins having a common
XX CC function to be identified. Identification of novel SH3 proteins will be
XX CC useful for a better understanding of cell growth, malignancy, signal
XX CC transduction processes, etc. New candidate drugs can be identified, and
XX CC their specificities (e.g. pharmacological activities) can be assessed
XX CC using the method of the invention
XX SQ Sequence 17 AA;
XX Query Match 100.0%; Score 37; DB 2; Length 17;
XX Best Local Similarity 100.0%; Pred. No. 36;
XX Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX QY 1 ILAPPVP 7
XX Db |||||
XX 6 ILAPPVP 12
XX RESULT 15
XX AA016938
XX ID AA016938 standard; peptide; 30 AA.
XX AC AA016938;
XX XX
XX DT 27-JUN-1997 (first entry)
XX XX
XX DE Random recombinant SH3 domain binding peptide.
XX XX

```

KW Src; SH3; Src homology region 3; binding affinity; oncogenic protein;
 KW protein tyrosine kinase; signal transduction; RNA processing;
 XX trafficking; translation.
 OS Synthetic.
 XX WO9603649-A1.
 PN XX
 PD 08-FEB-1996.
 XX
 PF 24-JUL-1995; 95WO-US009382.
 XX
 PR 22-JUL-1994; 94US-00278865.
 PR 07-JUN-1995; 95US-00483555.
 XX
 PA (UYNC-) UNIV NORTH CAROLINA.
 XX
 PI Sparks AB, Kay BK, Thorn JM, Quilliam LA, Der CJ;
 XX WPI; 1996-117151/12.
 DR
 PT Peptide with binding affinity for Src homology region 3 (SH3) domains of
 PT proteins - useful for e.g. modulating signal transduction pathways at the
 PT cellular level, esp. protein tyrosine kinase-mediated.
 XX
 PS Disclosure; Fig 1; 116pp; English.
 XX
 CC AAW16924-W16948 are random recombinant peptides derived from one of three
 CC peptide libraries, T9, T12 and R8C. The peptides are all SH3 domain-
 CC binding peptides. SH3 binding peptides are useful in modulating signal
 CC transduction pathways at the cellular level (especially protein tyrosine
 CC kinase-mediated), modulating oncogenic protein activity, or providing
 CC compounds for the development of drugs with the ability to modulate broad
 CC classes, as well as specific classes, of proteins involved in signal
 CC transduction and also for regulating the processing, trafficking or
 CC translation of RNA. Conjugates of the peptides with detectable labels or
 CC imaging agents are useful for imaging cells, tissues and organs in which
 CC Src or Src-related proteins are expressed
 XX
 SQ Sequence 30 AA;
 Query Match 100.0%; Score 37; DB 2; Length 30;
 Best Local Similarity 100.0%; Pred. No. 62;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ILAPPVP 7
 Db 16 ILAPPVP 22
 RESULT 16
 AAW25501
 ID AAW25501 standard; peptide; 30 AA.
 XX
 AC AAW25501;
 XX
 DT 27-MAR-1998 (first entry)
 XX
 DE Random peptide recombinant clone T12.SRC3.1.
 XX
 KW Cortactin; SH3 domain; binding peptide; Src homology region 3;
 KW tyrosine kinase; immune response; lymphokine; interleukin 1; Nck; Abl;
 KW PLCgamma; p53bp2; Crk; Yes; Grb2.
 XX
 OS Synthetic.
 OS Unidentified.
 XX
 PN WO9730074-A1.
 XX
 PD 21-AUG-1997.
 XX
 PF 14-FEB-1997; 97WO-US002298.
 XX

PR 16-FEB-1996; 96US-00602999.
 XX
 PA (CYTO-) CYTOGEN CORP.
 PA (UYNC-) UNIV NORTH CAROLINA.
 XX
 PI Sparks AB, Kay BK, Thorn JM, Quilliam LA, Der CJ, Fowlkes DM;
 PI Rider JE;
 XX
 DR WPI; 1997-424972/39.
 XX
 PT Src homology region 3 binding peptide - used to activate Src tyrosine
 PT kinase(s) and to stimulate immune response by increasing production of
 PT certain lymphokine(s), e.g. interleukin-1.
 XX
 PS Disclosure; Fig 5; 131pp; English.
 XX
 CC The present sequence represents a random peptide recombinant isolated by
 CC the method of the present invention. SH3 (Src homology region 3) binding
 CC peptides are selected from: (a) peptides which bind the SH3 domain of
 CC Cortactin; (b) peptides which bind the middle SH3 domain of Nck; (c)
 CC peptides which bind the SH3 domain of Abl; (d) peptides which bind the
 CC SH3 domain of Src; (e) peptides which bind the SH3 domain of PLC gamma;
 CC (f) peptides which bind the SH3 domain of p53bp2; (g) peptides which bind
 CC the amino-terminal SH3 domain of Crk; (h) peptides which bind the SH3
 CC domain of Yes; and (i) peptides which bind the amino-terminal SH3 domain
 CC of Grb2. The purified binding peptides can be used in the method to
 CC identify inhibitors of their binding to their respective SH3 domains,
 CC which could be used to modulate the pharmacological activity of proteins
 CC or polypeptide containing the SH3 domain. The peptides can also be used
 CC to activate Src or Src-related protein tyrosine kinases, to stimulate the
 CC immune response by increasing the production of certain lymphokines, e.g.
 CC tumour necrosis factor-alpha and interleukin-1, or to deliver a
 CC conjugated molecule to certain cellular compartments containing Src or
 CC Src related proteins
 XX
 SQ Sequence 30 AA;
 Query Match 100.0%; Score 37; DB 2; Length 30;
 Best Local Similarity 100.0%; Pred. No. 62;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ILAPPVP 7
 Db 16 ILAPPVP 22
 RESULT 17
 ABB62726
 ID ABB62726 standard; protein; 1767 AA.
 XX
 AC ABB62726;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster polypeptide SEQ ID NO 14970.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 XX
 OS Drosophila melanogaster.
 XX
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US009231.
 XX
 PR 23-MAR-2000; 2000US-0191637P.
 PR 11-JUL-2000; 2000US-00614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.
 DR N-PSDB; ABL06829.
 XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions.
 XX Disclosure; SEQ ID NO 14970; 21pp + Sequence Listing; English.
 XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
 CC ABB72072). The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX Sequence 1767 AA;
 SQ
 Query Match 100.0%; Score 37; DB 4; Length 1767;
 Best Local Similarity 100.0%; Pred. No. 2.8e+03;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ILAPPVP 7
 Db 1101 ILAPPVP 1107
 RESULT 18
 AAW25387
 ID AAW25387 standard; peptide; 16 AA.
 AC AAW25387;
 XX
 XX 27-MAR-1998 (first entry)
 DT
 DE PLC gamma SH3 domain binding peptide SEQ ID NO:178.
 XX
 XX Cortactin; SH3 domain; binding peptide; Src homology region 3;
 KW tyrosine kinase; immune response; lymphokine; interleukin 1; Nck; Abl;
 KW PLCgamma; p53bp2; Crk; Yes; Grb2.
 XX Synthetic.
 OS Unidentified.
 OS
 XX WO9730074-A1.
 PN
 XX 21-AUG-1997.
 PD
 XX 14-FEB-1997; 97WO-US002298.
 PF
 XX 16-FEB-1996; 96US-00602999.
 PR
 XX (CYTO-) CYTOGEN CORP.
 PA (UYNC-) UNIV NORTH CAROLINA.
 PA
 XX Sparks AB, Kay BK, Thorn JM, Quilliam LA, Der CJ, Fowlkes DM;
 PI Rider JE;
 PI WPI; 1997-424972/39.
 DR
 XX Src homology region 3 binding peptide - used to activate Src tyrosine
 PT kinase(s) and to stimulate immune response by increasing production of
 PT certain lymphokine(s), e.g. interleukin-1.
 XX
 XX Claim 15; Page 98; 131pp; English.
 PS
 XX The present sequence represents a Src homology region 3 (SH3) binding
 CC peptide. SH3 binding peptides are selected from: (a) peptides which bind

CC the SH3 domain of Cortactin; (b) peptides which bind the middle SH3
 CC domain of Nck; (c) peptides which bind the SH3 domain of Abl; (d)
 CC peptides which bind the SH3 domain of Src; (e) peptides which bind the
 CC SH3 domain of PLC gamma; (f) peptides which bind the SH3 domain of p53bp2
 CC ; (g) peptides which bind the amino-terminal SH3 domain of Crk; (h)
 CC peptides which bind the SH3 domain of Yes; and (i) peptides which bind
 CC the amino-terminal SH3 domain of Grb2. The purified binding peptides can
 CC be used in the method to identify inhibitors of their binding to their
 CC respective SH3 domains, which could be used to modulate the
 CC pharmacological activity of proteins or polypeptide containing the SH3
 CC domain. The peptides can also be used to activate Src or Src-related
 CC protein tyrosine kinases, to stimulate the immune response by increasing
 CC the production of certain lymphokines, e.g. tumour necrosis factor-alpha
 CC and interleukin-1, or to deliver a conjugated molecule to certain
 CC cellular compartments containing Src or Src related proteins
 XX
 SQ Sequence 16 AA;
 Query Match 97.3%; Score 36; DB 2; Length 16;
 Best Local Similarity 85.7%; Pred. No. 49;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ILAPPVP 7
 Db 1 VLAPPVP 7
 RESULT 19
 ABP43530
 ID ABP43530 standard; protein; 137 AA.
 XX
 AC ABP43530;
 XX
 XX 08-AUG-2002 (first entry)
 DT
 DE Human secreted protein (SCEP) 54.
 XX
 XX Human; secreted protein; SECP; SSCP expression; gene therapy;
 KW protein therapy; Immune system disorders; AIDS; thymic hypoplasia;
 KW anaemia; asthma; Crohn's disease; neurological disorder; epilepsy;
 KW Huntington's disease; dementia; Parkinson's disease; Down's syndrome;
 KW developmental disorder; cell proliferative disorder; cancer.
 XX Homo sapiens.
 OS
 XX WO200226982-A2.
 PN
 XX 04-APR-2002.
 PD
 XX 25-SEP-2001; 2001WO-US030042.
 PF
 XX 29-SEP-2000; 2000US-0236869P.
 PR 11-OCT-2000; 2000US-0239812P.
 PR 12-OCT-2000; 2000US-0240108P.
 PR 17-OCT-2000; 2000US-0241282P.
 PR 20-OCT-2000; 2000US-0242218P.
 XX
 XX (INCY-) INCYTE GENOMICS INC.
 XX
 XX Yue H, Tang YT, Nguyen DB, Yao MG, Xu Y, Tribouley CM;
 PI Sarjanwala MS, Wallia NK, Baughn MR, Sapperstein SK, Lal P;
 PI Thornton W, Gandhi AR, Ramkumar J, Elliott VS, Arvizu C;
 PI Thangavelu K, Gietzen KJ, Ding L, Au-Young J, Tran B, Policky JL;
 PI Lee S, Lu DAM, Burford N, Warren BA, Gururajan R, Duggan BM;
 PI Honchell CD, Hafalia AJA;
 XX WPI: 2002-394239/42.
 DR N-PSDB; AEN99413.
 XX
 XX New human secreted proteins, useful for diagnosing, treating or
 PT preventing immune system disorders (e.g. Crohn's disease), neurological
 PT disorders (e.g. Parkinson's disease), or cell proliferative disorders
 PT (e.g. cancers).

XX Claim 1; Page 186; 238pp; English.

PS The invention comprises the amino acid and coding sequences of human

CC secreted proteins (SECP). The SECP DNA and amino acid sequences of the

CC invention are useful for treating/preventing disorders associated with

CC decreased or elevated expression of SECP. The SECP DNA and protein

CC sequences are specifically useful for treating/preventing (i.e. gene

CC therapy and protein therapy): immune system disorders (e.g. AIDS, thymic

CC hypoplasia, anaemia, asthma or Crohn's disease); neurological disorders

CC (e.g. epilepsy, Huntington's disease, dementia or Parkinson's disease);

CC developmental disorders (e.g. Down's syndrome); and cell proliferative

CC disorders (e.g. cancer). The proteins ABP43477 - ABP43543 represent the

CC human secreted proteins (SECP) of the invention

XX

SQ Sequence 137 AA;

Query Match 97.3%; Score 36; DB 5; Length 137;

Best Local Similarity 85.7%; Pred. No. 3.7e+02;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAPPVP 7

Db 15 VLAPPVP 21

RESULT 20

AAG83323

ID AAG83323 standard; protein; 159 AA.

AC AAG83323;

XX

DT 06-SEP-2001 (first entry)

XX

DE P patens lipid metabolism related protein #41.

XX

KW Moss; LMRP; lipid metabolism related protein; polyunsaturated fatty acid;

KW fine chemical; transgenic plant.

XX

OS Physcomitrella patens.

XX

PI Lerchl J, Renz A, Ehrhardt T, Reindl A, Cirpus P, Bischoff P;

XX Frank M, Freund A, Duwenig E, Schmidt R, Reski R;

PN WO200138541-A1.

XX

PD 31-MAY-2001.

XX

PF 25-NOV-1999; 99WO-EP009108.

XX

PR 25-NOV-1999; 99WO-EP009108.

XX

XX (BADI) BASF PLANT SCI GMBH.

PA

XX Lerchl J, Renz A, Ehrhardt T, Reindl A, Cirpus P, Bischoff P;

PI Frank M, Freund A, Duwenig E, Schmidt R, Reski R;

XX WPI; 2001-381293/40.

XX

XX New isolated nucleic acid molecule encoding Lipid Metabolism Related

PT Proteins useful in the production of fine chemicals.

XX

PS Claim 31; Page 103; 113pp; English.

XX

CC The present invention provides the protein and coding sequences of a

CC number of moss lipid metabolism related proteins (LMRPs). The moss

CC Physcomitrella patens is one of the few plants able to produce

CC polyunsaturated fatty acids, and the sequences can be used to create

CC transgenic plants also capable of producing them. They can also be used

CC to identify the presence of P. patens and in the production of fine

CC chemicals. The present sequence is one of the proteins of the invention

XX

SQ Sequence 159 AA;

Query Match 97.3%; Score 36; DB 4; Length 159;

Best Local Similarity 85.7%; Pred. No. 4.2e+02;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAPPVP 7

Db 41 VLAPPVP 47

RESULT 21

AAG80875

ID AAG80875 standard; protein; 159 AA.

XX

AC AAG80875;

XX

DT 28-AUG-2001 (first entry)

XX

DE Lipid biosynthesis protein sequence #33.

XX

KW Moss; Physcomitrella patens; lipid metabolism related protein; LMRP;

KW lipid biosynthesis; lipid modification; lipid degradation; cofactor;

KW fatty acid transport; genetic engineering; fatty acid; enzyme; plant;

KW microorganism; polyunsaturated fatty acid; oilseed plant; maize; wheat;

KW biotic stress tolerance; abiotic stress tolerance; rye; oat; triticale;

KW rice; barley; soybean; peanut; cotton; rapeseed; canola; manihot; pepper;

KW sunflower; tagetes; potato; tobacco; eggplant; tomato; Vicia; pea;

KW alfalfa; coffee; cacao; tea; Salix; oil palm; coconut; perennial grass;

KW forage crop.

XX

OS Physcomitrella patens.

XX

PN WO200138484-A2.

XX

PD 31-MAY-2001.

XX

PF 22-NOV-2000; 2000WO-EP011615.

XX

PR 25-NOV-1999; 99WO-EP009108.

XX

XX (BADI) BASF PLANT SCI GMBH.

PA

XX Lerchl J, Renz A, Ehrhardt T, Reindl A, Cirpus P, Bischoff P;

PI Frank M, Freund A, Duwenig E, Schmidt R, Reski R;

XX WPI; 2001-367669/38.

XX

XX Nucleic acids encoding lipid metabolism related proteins from

PT Physcomitrella patens useful to produce fine chemicals in modified

PT organisms, particularly polyunsaturated fatty acids in oilseed plants.

XX

PS Claim 31; Page 113; 120pp; English.

XX

CC The present invention describes isolated nucleic acid sequences which

CC encode lipid metabolism related proteins (LMRP). The LMRP nucleic acids

CC can be used to modify lipids and fatty acids, cofactors and enzymes in

CC microorganisms and plants, particularly to produce polyunsaturated fatty

CC acids, and are especially useful in oilseed plants. The nucleic acids may

CC also confer biotic or abiotic stress tolerance, particularly to maize,

CC wheat, rye, oat, triticale, rice, barley, soybean, peanut, cotton,

CC rapeseed, canola, manihot, pepper, sunflower, tagetes, potato, tobacco,

CC eggplant, tomato, Vicia species, pea, alfalfa, coffee, cacao, tea, Salix

CC species, oil palm, coconut, perennial grasses and forage crops. AAH50878

CC to AAH50892 represent primers used in the exemplification of the present

CC invention. AAH50883 to AAH50968 represents LMRP nucleotide sequences, and

CC AAG80843 to AAG80928 represent LMRP protein sequences, given in the

CC present invention

XX

SQ Sequence 159 AA;

Query Match 97.3%; Score 36; DB 4; Length 159;

Best Local Similarity 85.7%; Pred. No. 4.2e+02;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAPPVP 7

Db 41 VLAPPVP 47

Db 41 VLAPPVP 47

RESULT 22
ABB50640
ID ABB50640 standard; protein; 200 AA.

XX
AC ABB50640;
XX
DT 07-FEB-2002 (first entry)
XX
DE Human secreted protein encoded by gene 39 SEQ ID NO:588.

XX
KW Human; secreted protein; immunomodulatory; antisclerotic; anti-HIV;
KW dermatological; immunosuppressive; antiinflammatory; immunostimulant;
KW cytostatic; cardiac; anti-angiogenic; ophthalmological;
KW neuroprotective; nootropic; anticonvulsant; antialzheimers; vulnerary;
KW antiparkinsonian; antimicrobial; gene therapy; vaccine; immune disorder;
KW multiple sclerosis; systemic lupus erythematosus; HIV infection; cancer;
KW human immunodeficiency virus; hyperproliferative disorder; wound healing;
KW Gaucher's disease; cardiovascular disease; Scimitar syndrome; chemotaxis;
KW Chaga's cardiomyopathy; coronary arteriosclerosis; angiogenic disorder;
KW corneal graft neovascularisation; diabetic retinopathy; regeneration;
KW neurological disorder; Huntington's chorea; Alzheimer's disease;
KW Parkinson's disease; infectious disease.

XX
OS Homo sapiens.

XX
PN WO200162891-A2.

XX
PD 30-AUG-2001.

XX
PF 21-FEB-2001; 2001WO-US005614.

XX
PR 24-FEB-2000; 2000US-0184836P.

XX
PR 29-MAR-2000; 2000US-0193170P.

XX
PR (HUMA-) HUMAN GENOME SCI INC.

XX
PI Ni J, Ebner R, Lafleur DM, Moore PA, Olsen HS, Rosen CA;
PI Ruben SM, Soppet DR, Young PE, Shi Y, Florence KA, Wei Y;
PI Florence C, Hu J, Li Y, Kyaw H, Fischer CL, Ferrie AM, Fan P;
PI Feng P, Endress GA, Dillon PJ, Carter KC, Brewer LA, Yu G, Zeng Z;
PI Greene JM;

XX
XX WPI; 2001-625724/72.

XX
XX Nucleic acids encoding 207 human secreted polypeptides, useful for
XX preventing, diagnosing and/or treating, e.g. cancers, Parkinson's disease
XX and diabetic retinopathy.

XX
XX Disclosure; Page 90; 1533pp; English.

XX
XX ABB50301 to ABB51287 and ABA83194 to ABA83441 represent human secreted
XX proteins (I) and polynucleotide (II) sequences. (I) and (II) have various
XX activities based on the tissues and cells the genes are expressed in.
XX Example of these activities include: immunomodulatory; antisclerotic;
XX dermatological; immunosuppressive; antiinflammatory; immunostimulant;
XX anti-HIV; cytostatic; cardiac; anti-angiogenic; ophthalmological;
XX neuroprotective; nootropic; anticonvulsant; antialzheimers; vascular;
XX antiparkinsonian; antimicrobial; and vulnerary. (I) and (II) can be used
XX in gene therapy and vaccine production. (I) and (II) can be used in the
XX prevention, diagnosis and treatment of immune disorders (e.g. multiple
XX sclerosis, systemic lupus erythematosus and human immunodeficiency virus
XX (HIV) infections), hyperproliferative disorders (e.g. cancers and
XX Gaucher's disease), cardiovascular diseases (e.g. Scimitar syndrome,
XX Chaga's cardiomyopathy and coronary arteriosclerosis), angiogenic
XX disorders (e.g. corneal graft neovascularisation and diabetic
XX retinopathy), neurological disorders (e.g. Huntington's chorea,
XX Alzheimer's disease and Parkinson's disease), infectious diseases and/or
XX for promoting wound healing, regeneration and/or chemotaxis. ABA83185 to
XX ABA83193 and ABB50300 represent sequences used in the exemplification of
XX the present invention.

XX
SQ Sequence 200 AA;

Query Match 97.3%; Score 36; DB 4; Length 200;
Best Local Similarity 85.7%; Pred. NO. 5.2e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAPPVP 7
Db 160 VLAPPVP 166
:|||||

RESULT 23
ABO44897
ID ABO44897 standard; protein; 200 AA.

XX
AC ABO44897;
XX
DT 02-OCT-2003 (first entry)
XX
DE Novel human secreted protein #39 fragment #3.

XX
KW Human; gene therapy; autoimmune disorder; multiple sclerosis; cancer;
KW systemic lupus erythematosus; haematopoietic cell disorder; allergy;
KW agammaglobulinaemia; ataxia telangiectasia; blood coagulation disorder;
KW afibrinogenaemia; thrombocytopenia; graft-versus-host disease; arthritis;
KW inflammatory condition; ischaemia-reperfusion injury; infectious disease;
KW hyperproliferative disorder; purpura; viral infection; regeneration;
KW bacterial infection; ulcer; Alzheimer's disease.

XX
OS Homo sapiens.

XX
PN US2003065160-A1.

XX
PD 03-APR-2003.

XX
PF 07-DEC-2001; 2001US-00004860.

XX
XX 06-JUN-1997; 97US-0048875P.
XX 06-JUN-1997; 97US-0048876P.
XX 06-JUN-1997; 97US-0048877P.
XX 06-JUN-1997; 97US-0048878P.
XX 06-JUN-1997; 97US-0048880P.
XX 06-JUN-1997; 97US-0048881P.
XX 06-JUN-1997; 97US-0048882P.
XX 06-JUN-1997; 97US-0048883P.
XX 06-JUN-1997; 97US-0048884P.
XX 06-JUN-1997; 97US-0048885P.
XX 06-JUN-1997; 97US-0048886P.
XX 06-JUN-1997; 97US-0048887P.
XX 06-JUN-1997; 97US-0048888P.
XX 06-JUN-1997; 97US-0048889P.
XX 06-JUN-1997; 97US-0048890P.
XX 06-JUN-1997; 97US-0048891P.
XX 06-JUN-1997; 97US-00488916P.
XX 06-JUN-1997; 97US-0048917P.
XX 06-JUN-1997; 97US-0048949P.
XX 06-JUN-1997; 97US-0048962P.
XX 06-JUN-1997; 97US-0048963P.
XX 06-JUN-1997; 97US-0048964P.
XX 06-JUN-1997; 97US-0048970P.
XX 06-JUN-1997; 97US-0048971P.
XX 06-JUN-1997; 97US-0048972P.
XX 06-JUN-1997; 97US-0048974P.
XX 06-JUN-1997; 97US-0049019P.
XX 06-JUN-1997; 97US-0049020P.
XX 06-JUN-1997; 97US-0049373P.
XX 06-JUN-1997; 97US-0049374P.

PR	06-JUN-1997;	97US-0049375P.	XX	Sequence 200 AA;	
PR	05-SEP-1997;	97US-0057584P.	SQ		
PR	05-SEP-1997;	97US-0057622P.		Query Match	97.3%; Score 36; DB 6; Length 200;
PR	05-SEP-1997;	97US-0057628P.		Best Local Similarity	85.7%; Pred. No. 5.2e+02;
PR	05-SEP-1997;	97US-0057629P.		Matches	6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
PR	05-SEP-1997;	97US-0057634P.			
PR	05-SEP-1997;	97US-0057635P.			
PR	05-SEP-1997;	97US-0057642P.	QY	1 ILAPPVP 7	
PR	05-SEP-1997;	97US-0057643P.			
PR	05-SEP-1997;	97US-0057644P.	Db	160 VLAPPVP 166	
PR	05-SEP-1997;	97US-0057645P.			
PR	05-SEP-1997;	97US-0057646P.			
PR	05-SEP-1997;	97US-0057647P.			
PR	05-SEP-1997;	97US-0057648P.			
PR	05-SEP-1997;	97US-0057649P.			
PR	05-SEP-1997;	97US-0057650P.			
PR	05-SEP-1997;	97US-0057651P.			
PR	05-SEP-1997;	97US-0057654P.			
PR	05-SEP-1997;	97US-0057661P.			
PR	05-SEP-1997;	97US-0057662P.			
PR	05-SEP-1997;	97US-0057666P.			
PR	05-SEP-1997;	97US-0057667P.			
PR	05-SEP-1997;	97US-0057668P.			
PR	05-SEP-1997;	97US-0057760P.			
PR	05-SEP-1997;	97US-0057761P.			
PR	05-SEP-1997;	97US-0057762P.			
PR	05-SEP-1997;	97US-0057763P.			
PR	05-SEP-1997;	97US-0057764P.			
PR	05-SEP-1997;	97US-0057765P.			
PR	05-SEP-1997;	97US-0057769P.			
PR	05-SEP-1997;	97US-0057770P.			
PR	05-SEP-1997;	97US-0057771P.			
PR	05-SEP-1997;	97US-0057774P.			
PR	05-SEP-1997;	97US-0057775P.			
PR	05-SEP-1997;	97US-0057776P.			
PR	05-SEP-1997;	97US-0057777P.			
PR	05-SEP-1997;	97US-0057778P.			
PR	18-DEC-1997;	97US-0070923P.			
PR	04-JUN-1998;	98WO-US011422.			
PR	15-JUL-1998;	98US-0092921P.			
PR	30-JUL-1998;	98US-0094657P.			
PR	04-DEC-1998;	98US-00205258.			
PA	(HUMA-) HUMAN GENOME SCI INC.				
XX	Young P, Greene JM, Ferrie AM, Ruben SM, Rosen CA, Hu J;				
PI	Olsen HS, Ebner R, Brewer LA, Moore PA, Shi Y, Florence C;				
PI	Florence K, Laflaur DW, Ni J, Pan P, Wei Y, Fischer CL, Soppet DR;				
PI	Li Y, Zeng Z, Kyaw H, Yu G, Feng P, Dillon PJ, Endress GA,				
PI	Carter KC;				
XX	WPI; 2003-540804/51.				
DR					
XX					
XX	New isolated protein, useful for preparing a composition for diagnosing				
PT	or treating cancer, inflammatory, immune or infectious diseases.				
PT					
XX					
PS	Disclosure; Page 23; 172pp; English.				
XX					
CC	The invention relates to an isolated HEMAF80 protein. The protein is				
CC	useful for preparing a composition for diagnosing or treating autoimmune				
CC	disorders e.g. multiple sclerosis and systemic lupus erythematosus;				
CC	haematopoietic cell disorders e.g. agammaglobulinaemia and ataxia				
CC	relangiectasia; blood coagulation disorders e.g. afibrinogenaemia and				
CC	thrombocytopenia; allergy; graft-versus-host disease; inflammatory				
CC	conditions e.g. ischaemia-reperfusion injury and arthritis;				
CC	hyperproliferative disorders e.g. cancer and purpura; infectious disease				
CC	e.g. viral infection and bacterial infection. The polynucleotide or				
CC	protein can be used to regenerate damaged tissue e.g. ulcers and				
CC	Alzheimer's disease. The present sequence represents the amino acid				
CC	sequence of a novel human secreted protein fragment. Note: The sequence				
CC	data for this patent did not form part of the printed specification but				
CC	was obtained in electronic format directly from USPTO at				
CC	seqdata.uspto.gov/sequence.html?DocID=20030065160				

XX	Secreted protein; precerebellin-like protein; sepsis; acne; psoriasis;				
KW	neurodegenerative disorder; behavioural disorder; Alzheimer's disease;				
KW	Parkinson's disease; Huntington's disease; schizophrenia; mania;				
KW	dementia; paranoia; psychosis; autism; immune disorder; infection;				
KW	inflammation; allergy; liver disorder; hepatoblastoma; jaundice;				
KW	hepatitis; immunological disorder; AIDS; leukaemia; rheumatoid arthritis;				
KW	cancer.				
XX	Unidentified.				
XX	US6525174-B1.				
XX	25-FEB-2003.				
XX	04-DEC-1998;	98US-00205258.			
XX	06-JUN-1997;	97US-0048875P.			
PR	06-JUN-1997;	97US-0048876P.			
PR	06-JUN-1997;	97US-0048877P.			
PR	06-JUN-1997;	97US-0048878P.			
PR	06-JUN-1997;	97US-0048880P.			
PR	06-JUN-1997;	97US-0048881P.			
PR	06-JUN-1997;	97US-0048882P.			
PR	06-JUN-1997;	97US-0048883P.			
PR	06-JUN-1997;	97US-0048884P.			
PR	06-JUN-1997;	97US-0048885P.			
PR	06-JUN-1997;	97US-0048892P.			
PR	06-JUN-1997;	97US-0048893P.			
PR	06-JUN-1997;	97US-0048894P.			
PR	06-JUN-1997;	97US-0048895P.			
PR	06-JUN-1997;	97US-0048896P.			
PR	06-JUN-1997;	97US-0048897P.			
PR	06-JUN-1997;	97US-0048898P.			
PR	06-JUN-1997;	97US-0048899P.			
PR	06-JUN-1997;	97US-0048900P.			
PR	06-JUN-1997;	97US-0048901P.			
PR	06-JUN-1997;	97US-0048915P.			
PR	06-JUN-1997;	97US-0048916P.			
PR	06-JUN-1997;	97US-0048917P.			
PR	06-JUN-1997;	97US-0048949P.			
PR	06-JUN-1997;	97US-0048962P.			
PR	06-JUN-1997;	97US-0048963P.			
PR	06-JUN-1997;	97US-0048964P.			
PR	06-JUN-1997;	97US-0048970P.			
PR	06-JUN-1997;	97US-0048971P.			
PR	06-JUN-1997;	97US-0048972P.			
PR	06-JUN-1997;	97US-0048974P.			
PR	06-JUN-1997;	97US-0049019P.			
PR	06-JUN-1997;	97US-0049020P.			
PR	06-JUN-1997;	97US-0049373P.			
PR	06-JUN-1997;	97US-0049374P.			
PR	06-JUN-1997;	97US-0049377P.			

AAU17916
ID AAU17916 standard; protein; 116 AA.
XX AC AAU17916;
XX DT 07-NOV-2001 (first entry)
XX DE Novel human respiratory antigen #232.
XX KW Human; respiratory antigen; respiratory disorder; throat disorder;
KW lung disorder; nose disorder; lung cancer; gene therapy; cytostatic;
KW anti allergic; anti asthmatic; anti inflammatory; olfactory;
KW respiratory active.
XX OS Homo sapiens.
XX PN WO200155448-A1.
XX PD 02-AUG-2001.
XX PF 17-JAN-2001; 2001WO-US001333.
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 13-OCT-2000; 2000US-0239933P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246521P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246603P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.

PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 06-DEC-2000; 2000US-0256719P.
 PR 08-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 08-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Rosen CA, Barash SC, Ruben SM;
 XX
 XX WPI; 2001-476224/51.
 DR N-PSDB; AAS28100.
 XX
 XX Isolated polypeptide for treating, preventing and/ or prognosing
 PT disorders related to the respiratory system including respiratory cancers
 PT and also for testing and detection e.g. diagnosis.
 XX
 XX Claim 11; SED ID No 534; 546pp; English.
 XX
 XX The present invention relates to the isolation of novel human respiratory
 CC antigens, and cDNA (AAS27869-AAS28159) and genomic sequences encoding for
 CC these polypeptides. The sequences of the invention are useful for
 CC preventing, treating and/or prognosing disorders related to the
 CC respiratory system including throat disorders (e.g. vocal cord paralysis,
 CC tonsillitis, and laryngitis), lung disorders e.g. pneumonia, allergic
 CC disorders e.g. asthma, pleurisy, cystic fibrosis, emphysema, nose
 CC disorders and cancers of the respiratory tissues e.g. lung cancer. The
 CC polynucleotide sequences of the invention are useful in gene therapy and
 CC antigenase therapy. AAU17685-AAU17975 represent novel human respiratory
 CC antigens. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 XX Sequence 116 AA;
 SQ
 Query Match 94.6%; Score 35; DB 4; Length 116;
 Best Local Similarity 85.7%; Pred. No. 4.5e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ILAPPVP 7
 Db :|||||
 77 LLAPPVP 83
 RESULT 27
 ADG41296
 ID ADG41296 standard; protein; 116 AA.
 XX
 AC ADG41296;
 XX
 XX 26-FEB-2004 (first entry)
 DT
 XX Human respiratory system associated protein seq id 534.
 DE
 XX antinflammatory; antiallergic; antiaesthetic; cytostatic; gene therapy;
 XX respiratory system antigen;
 KW human respiratory system associated polynucleotide;
 KW human respiratory system associated polynucleotide;
 KW tonsillitis; laryngitis; lung disorder; pneumonia; allergic disorder;
 KW asthma; eosinophilic pneumonia; pleurisy; cystic fibrosis; emphysema;
 KW histiocytosis; sarcoidosis; nose disorder; rhinitis; sinusitis; neoplasm;
 KW cancer; respiratory tissue cancer; throat cancer; lung cancer;
 KW cancer of the nose; gene therapy; chromosome identification; forensic;
 KW human respiratory system associated protein; human.
 XX
 XX Homo sapiens.

XX US2003215893-A1.
 XX 20-NOV-2003.
 XX
 XX 07-AUG-2002; 2002US-00212872.
 XX
 XX 31-JAN-2000; 2000US-0179065P.
 PR 04-FEB-2000; 2000US-0180628P.
 PR 24-FEB-2000; 2000US-0184664P.
 PR 02-MAR-2000; 2000US-0186350P.
 PR 16-MAR-2000; 2000US-0189874P.
 PR 17-MAR-2000; 2000US-0190076P.
 PR 18-APR-2000; 2000US-0198123P.
 PR 19-MAY-2000; 2000US-0205515P.
 PR 07-JUN-2000; 2000US-0209467P.
 PR 28-JUN-2000; 2000US-0214886P.
 PR 30-JUN-2000; 2000US-0215135P.
 PR 07-JUL-2000; 2000US-0216647P.
 PR 07-JUL-2000; 2000US-0216880P.
 PR 11-JUL-2000; 2000US-0217487P.
 PR 11-JUL-2000; 2000US-0217496P.
 PR 14-JUL-2000; 2000US-0218290P.
 PR 26-JUL-2000; 2000US-0220963P.
 PR 26-JUL-2000; 2000US-0220964P.
 PR 14-AUG-2000; 2000US-0224518P.
 PR 14-AUG-2000; 2000US-0224519P.
 PR 14-AUG-2000; 2000US-0225213P.
 PR 14-AUG-2000; 2000US-0225214P.
 PR 14-AUG-2000; 2000US-0225266P.
 PR 14-AUG-2000; 2000US-0225267P.
 PR 14-AUG-2000; 2000US-0225268P.
 PR 14-AUG-2000; 2000US-0225270P.
 PR 14-AUG-2000; 2000US-0225447P.
 PR 14-AUG-2000; 2000US-0225757P.
 PR 14-AUG-2000; 2000US-0225758P.
 PR 14-AUG-2000; 2000US-0225759P.
 PR 18-AUG-2000; 2000US-0226279P.
 PR 22-AUG-2000; 2000US-0226681P.
 PR 22-AUG-2000; 2000US-0226688P.
 PR 22-AUG-2000; 2000US-0227182P.
 PR 23-AUG-2000; 2000US-0227009P.
 PR 30-AUG-2000; 2000US-0228924P.
 PR 01-SEP-2000; 2000US-0229287P.
 PR 01-SEP-2000; 2000US-0229343P.
 PR 01-SEP-2000; 2000US-0229344P.
 PR 01-SEP-2000; 2000US-0229345P.
 PR 05-SEP-2000; 2000US-0229509P.
 PR 05-SEP-2000; 2000US-0229513P.
 PR 06-SEP-2000; 2000US-0230437P.
 PR 06-SEP-2000; 2000US-0230438P.
 PR 08-SEP-2000; 2000US-0231242P.
 PR 08-SEP-2000; 2000US-0231243P.
 PR 08-SEP-2000; 2000US-0231244P.
 PR 08-SEP-2000; 2000US-0231413P.
 PR 08-SEP-2000; 2000US-0231414P.
 PR 08-SEP-2000; 2000US-0232080P.
 PR 12-SEP-2000; 2000US-0232081P.
 PR 14-SEP-2000; 2000US-0232397P.
 PR 14-SEP-2000; 2000US-0232398P.
 PR 14-SEP-2000; 2000US-0232399P.
 PR 14-SEP-2000; 2000US-0232400P.
 PR 14-SEP-2000; 2000US-0232401P.
 PR 14-SEP-2000; 2000US-0233063P.
 PR 14-SEP-2000; 2000US-0233064P.
 PR 14-SEP-2000; 2000US-0233065P.
 PR 21-SEP-2000; 2000US-0234223P.
 PR 21-SEP-2000; 2000US-0234274P.
 PR 25-SEP-2000; 2000US-0234997P.
 PR 25-SEP-2000; 2000US-0234998P.
 PR 26-SEP-2000; 2000US-0235484P.
 PR 27-SEP-2000; 2000US-0235834P.

PR 27-SEP-2000; 2000US-0235836P.
PR 28-SEP-2000; 2000US-0235935P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-02370337P.
PR 02-OCT-2000; 2000US-02370338P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
PR 17-JAN-2001; 2001US-00764860.
PR 14-FEB-2002; 2002US-00074095.
XX

PA (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Ruben SM, Barash SC;
XX
XX WPI; 2003-902033/82.
DR N-PSDB; ADG41004.
XX
XX Novel respiratory system antigen and polynucleotides encoding the
PT polypeptides, useful for treating diagnosing, treating or preventing
PT tonsillitis, pneumonia, asthma and cystic fibrosis, emphysema, throat
PT cancer.
XX
XX Claim 11; SEQ ID NO 534; 236pp; English.
XX
CC The invention describes an isolated polypeptide (I) comprising an amino
CC acid sequence that is at least 90% identical to polypeptide fragment of
CC any one of 299 respiratory system antigen sequences (PS) and having
CC biological activity, polypeptide domain or epitope of PS, full-length
CC protein of PS, or variant, allelic variant or species homolog of PS. (I)
CC or a polynucleotide (II) encoding (I) is also useful for diagnosing a
CC pathological condition or a susceptibility to a pathological condition in
CC a subject which involves determining the presence or absence of mutation
CC in (II) or determining the presence or amount of expression of (I) in a
CC biological sample and diagnosing a pathological condition based on the
CC result. The human respiratory system associated polynucleotides, the
CC polypeptides encoded by them, and antibodies that immunospecifically bind
CC these polypeptides are useful in diagnosis, treatment, prevention and/or
CC prognosis of disorders of respiratory system such as throat disorders
CC (e.g., vocal cord paralysis, tonsillitis, and laryngitis), lung disorders
CC (e.g., pneumonia), allergic disorders, (e.g., asthma and eosinophilic
CC pneumonia), pleurisy, cystic fibrosis, emphysema, histiocytosis,
CC sarcoidosis, nose disorders (rhinitis and sinusitis), neoplasms and/or
CC cancers of respiratory tissues (e.g., throat cancer, lung cancer, and
CC cancer of the nose). The polynucleotides are useful in gene therapy

Query Match 94.6%; Score 35; DB 7; Length 116;
Best Local Similarity 85.7%; Pred. No. 4.5e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 ILAPVP 7
:|||||
Db 77 LLAPVP 83

RESULT 28
ADI97070
ID ADI97070 standard; protein; 116 AA.
XX
AC ADI97070;
XX
XX 04-NOV-2004 (first entry)
XX
XX Human respiratory system associated polypeptide SeqIDS34.
XX
KW respiratory system-related polypeptide; antiasthmatic; antibacterial;
KW antiinflammatory; cytostatic; antianaemic; antiallergic; gene therapy;
KW pneumonia; lung cancer; cystic fibrosis; asthma; sarcoidosis; rhinitis;
KW anaemia; leukaemia; inflammation; sinusitis;
KW chronic obstructive pulmonary disease; infectious disease; human.
XX
OS Homo sapiens.
XX
XX US2003077704-A1.
XX
XX 24-APR-2003.
XX
XX 14-FEB-2002; 2002US-00074095.
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR

PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0228279P.
PR 22-AUG-2000; 2000US-0228681P.
PR 22-AUG-2000; 2000US-0228688P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0228287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.

PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
PR 17-JAN-2001; 2001US-00764860.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben SM, Barash SC;
XX WPI; 2003-765403/72.
XX N-PSDB; ADI96778.
XX
XX New human respiratory system-related polypeptide and genes, useful for
PT treating, preventing or diagnosing e.g. pneumonia, lung cancer, cystic
PT fibrosis, asthma, sarcoidosis, rhinitis, leukemia, inflammations or
XX sinusitis.
PS Claim 11; SEQ ID NO 534; 202pp; English.

XX This invention is related to a novel isolated polypeptide, which
CC comprises a human respiratory system-related polypeptide, and the DNA
CC sequence which encodes it. The invention may be useful for the
CC development of compounds with an antiasthmatic, antibacterial,
CC antiinflammatory, cytostatic, antianaemic or antiallergic activity. In
CC addition, the sequences disclosed may be useful for gene therapy. The
CC polypeptide or polynucleotide is useful for treating, preventing or
CC ameliorating a medical condition, for example pneumonia, lung cancer,
CC cystic fibrosis, asthma, sarcoidosis, rhinitis, anaemia, leukaemia,
CC inflammations, sinusitis, chronic obstructive pulmonary disease or
CC infectious diseases. The polypeptide or polynucleotide is also useful for
CC diagnosing any of these diseases or a susceptibility to the disease. The
CC present sequence is that of a human respiratory system associated
CC polypeptide of the invention.

XX SQ Sequence 116 AA;
Query Match 94.6%; Score 35; DB 7; Length 116;
Best Local Similarity 85.7%; Pred. No. 4.5e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAPPVP 7
Db 77 LLAPPVP 83
:|||||

RESULT 29
AAG10732
ID AAG10732 standard; protein; 132 AA.
AC AAG10732;
XX
XX DT 17-OCT-2000 (first entry)
XX Arabidopsis thaliana protein fragment SEQ ID NO: 9171.
DE
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
XX Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX PD 06-SEP-2000.
XX
XX PF 25-FEB-2000; 2000EP-00301439.
XX
XX PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.

PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.

```
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147302P.
PR 06-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149360P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149282P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153750P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156598P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157533P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 18-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.

Query Match          94.6%; Score 35; DB 3; Length 132;
Best Local Similarity 85.7%; Pred. No. 5.1e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAPPVP 7
Db 102 LLAPPVP 108

RESULT 30
ABM90479
ID ABM90479 standard; protein; 477 AA.
XX
AC ABM90479;
XX
DT 02-JUN-2005 (first entry)
XX
DE Rice abiotic stress responsive polypeptide SEQ ID NO:9201.
XX
KW abiotic stress tolerance; transgenic plant; plant; cereal; agriculture.
XX
OS Oryza sativa.
XX
PN WO2003008540-A2.
XX
PD 30-JAN-2003.
XX
PF 21-JUN-2002; 2002WO-US019668.
XX
PR 22-JUN-2001; 2001US-0300112P.
PR 24-AUG-2001; 2001US-0314662P.
PR 26-SEP-2001; 2001US-0325277P.
PR 21-NOV-2001; 2001US-0332132P.
XX
(PA (SYCN ) SYNGENTA PARTICIPATIONS AG.
XX
(Kraps J, Briggs SP, Cooper B, Glazebrook J, Goff SA, Katagiri F;
Moughamer T, Provart N, Ricke D, Zhu T;
WPI; 2003-248011/24.
XX
New stress-responsive nucleic acid, useful for altering the
responsiveness of a plant, e.g. cereal, to an abiotic stress such as cold
stress, salt stress or osmotic stress.
XX
Claim 1; SEQ ID NO 9201; 89pp; English.
XX
The invention relates to novel abiotic stress responsive polynucleotides
and polypeptides. Also disclosed are vectors, expression cassettes, host
cells, and plants containing such polynucleotides. Also disclosed are
methods for using the polynucleotides and polypeptides to alter the
responsiveness of a plant to abiotic stress. The invention is useful in
agriculture. The nucleic acid is useful for determining whether a test
plant has been exposed to an abiotic stress condition. It is also useful
for selecting an agent that alters abiotic stress regulated
polynucleotide expression in a plant cell, and to identify a homolog or
ortholog to an abiotic stress responsive polynucleotide. The nucleic acid
molecule and the polypeptide encoded by it are useful in altering the
responsiveness of a plant to an abiotic stress, such as cold stress, salt
stress, osmotic stress or any of their combinations. The present sequence
is used in the exemplification of the invention
XX
```

```
SQ Sequence 477 AA;
Query Match 94.6%; Score 35; DB 7; Length 477;
Best Local Similarity 85.7%; Pred. No. 1.7e+03;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAPPVP 7
Db 127 LLAPPVP 133
:|||||

RESULT 31
ABM87558
ID ABM87558 standard; protein; 477 AA.
XX
AC ABM87558;
XX
DT 02-JUN-2005 (first entry)
XX
DE Rice abiotic stress responsive polypeptide SEQ ID NO:5804.
XX
KW abiotic stress tolerance; transgenic plant; plant; cereal; agriculture.
XX
OS Oryza sativa.
XX
PN WO2003008540-A2.
XX
PD 30-JAN-2003.
XX
PF 21-JUN-2002; 2002WO-US019668.
XX
PR 22-JUN-2001; 2001US-0300112P.
XX
PR 24-AUG-2001; 2001US-0314662P.
XX
PR 26-SEP-2001; 2001US-0325277P.
XX
PR 21-NOV-2001; 2001US-0332132P.
XX
PA (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX
PI Kreps J, Briggs SP, Cooper B, Glazebrook J, Goff SA, Katagiri F;
PI Moughamer T, Provart N, Ricke D, Zhu T;
XX
DR WPI; 2003-248011/24.
XX
PT New stress-responsive nucleic acid, useful for altering the
PT responsiveness of a plant, e.g. cereal, to an abiotic stress such as cold
PT stress, salt stress or osmotic stress.
XX
PS Claim 1; SEQ ID NO 5804; 89pp; English.
XX
CC The invention relates to novel abiotic stress responsive polynucleotides
CC and polypeptides. Also disclosed are vectors, expression cassettes, host
CC cells, and plants containing such polynucleotides. Also disclosed are
CC methods for using the polynucleotides and polypeptides to alter the
CC responsiveness of a plant to abiotic stress. The invention is useful in
CC agriculture. The nucleic acid is useful for determining whether a test
CC plant has been exposed to an abiotic stress condition. It is also useful
CC for selecting an agent that alters abiotic stress regulated
CC polynucleotide expression in a plant cell, and to identify a homolog or
CC ortholog to an abiotic stress responsive polynucleotide. The nucleic acid
CC molecule and the polypeptide encoded by it are useful in altering the
CC responsiveness of a plant to an abiotic stress, such as cold stress, salt
CC stress, osmotic stress or any of their combinations. The present sequence
CC is used in the exemplification of the invention
XX
SQ Sequence 477 AA;
Query Match 94.6%; Score 35; DB 7; Length 477;
Best Local Similarity 85.7%; Pred. No. 1.7e+03;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAPPVP 7
Db 127 LLAPPVP 133
:|||||

RESULT 32
ADB65024
ID ADB65024 standard; protein; 113 AA.
XX
AC ADB65024;
XX
DT 04-DEC-2003 (first entry)
XX
DE Human protein encoded by clone SKNSH10001010.
XX
KW Human; pharmaceutical; diagnostic; gene therapy; tissue regeneration;
KW cell regeneration; membrane protein; signal transduction-related protein;
KW transcription-related protein; osteoporosis; neurological disease;
KW cancer; tumour.
XX
OS Homo sapiens.
XX
PN EP1308459-A2.
XX
PD 07-MAY-2003.
XX
PF 28-MAR-2002; 2002EP-00007401.
XX
PR 05-NOV-2001; 2001JP-00379298.
XX
PR 25-JAN-2002; 2002US-00350978.
XX
PA (HELI-) HELIX RES INST.
XX
PA (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;
XX
DR WPI; 2003-450961/43.
DR N-PSDB; ADB63054.
XX
PT New polynucleotides and polypeptides, useful for developing a diagnostic
PT marker or medicines for regulation of their expression and activity, or
PT as targets of gene therapy.
XX
PS Claim 1; Page; 222pp; English.
XX
CC The invention discloses a polynucleotide comprising a sequence selected
CC from 1970 fully defined nucleotide sequences which encode novel
CC polypeptides. Also claimed is a polypeptide encoded by the polynucleotide
CC or its partial peptide, an antibody binding to the polypeptide or peptide
CC of the polynucleotide, immunologically assaying the polypeptide or
CC peptide of the polynucleotide by contacting the polypeptide or peptide
CC with the antibody of the encoded protein, and observing the binding
CC between the two, a transformant carrying the polynucleotide in an
CC expressible manner and an antisense polynucleotide. The oligonucleotide
CC is useful as a primer for synthesising the polynucleotide, or as a probe
CC for detecting the polynucleotide. The polynucleotides and encoded
CC proteins are useful as pharmaceutical agents and many disease-related
CC genes may be included in them, for developing a diagnostic marker or
CC medicines for regulation of their expression and activity, or as targets
CC of gene therapy. The genes are involved in tissue and/or cell
CC regeneration. Membrane proteins, signal transduction-related proteins,
CC transcription-related proteins, disease-related proteins and genes
CC encoding them can be used as indicators for diseases (e.g. osteoporosis,
CC neurological diseases, cancer, tumours). The cDNA may be used to regulate
CC the activity or expression of the encoded protein to treat diseases. The
CC sequence presented is a protein of the invention. Note: Some of the
CC sequence data for this patent is not represented in the printed
CC specification, but is based on sequence information supplied by the
CC European Patent Office.
XX
SQ Sequence 113 AA;
Query Match 91.9%; Score 34; DB 7; Length 113;
Best Local Similarity 85.7%; Pred. No. 6.3e+02;
```

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAPVP 7
 ||:||||
 Db 70 ILSPPVP 76

RESULT 33
 ADC97354
 ID ADC97354 standard; protein; 177 AA.

XX AC ADC97354;

DT 01-JAN-2004 (first entry)

XX DE E. faecium protein sequence SEQ ID 6981.

XX KW Vaccine; urinary tract infection; bacteraemia; endocarditis; wound;
 KW abdominal-pelvic infection.

XX OS Enterococcus faecium.

XX PN US6583275-B1.

XX PD 24-JUN-2003.

XX PF 30-JUN-1998; 98US-00107532.

XX PR 02-JUL-1997; 97US-0051571P.

XX PR 14-MAY-1998; 98US-0085598P.

XX PA (GENO-) GENOME THERAPEUTICS CORP.

XX PI Doucette-Stamm LA, Bush D;

XX DR WPI; 2003-799836/75.

XX DR N-PSDB; ADC93700.

XX PT New isolated nucleic acid derived from Enterococcus faecium encoding an
 PT Enterococcus faecium polypeptide useful for detection, prevention and
 PT treatment of a pathological condition resulting from a bacterial
 PT infection.

XX PS Example 1; SEQ ID NO 6981; 243pp; English.

XX CC The invention relates to an isolated nucleic acid derived from
 CC Enterococcus faecium encoding an Enterococcus faecium polypeptide having
 CC one of 10 fully defined sequences given in the (or comprising 40
 CC sequential nucleotides chosen from any of the nucleic acids, its
 CC complement or sequences hybridising to it). Also included are a
 CC recombinant vector comprising the nucleic acid operably linked to
 CC transcription regulatory element, a cell comprising the vector and a
 CC single-stranded probe comprising the nucleic acid. The nucleic acids are
 CC chosen from 3654 disclosed sequences encoding 3654 disclosed proteins.
 CC The nucleic acids is useful for diagnosing pathological conditions
 CC resulting from E. faecium bacterial infection (e.g. urinary tract
 CC infection, bacteraemia, endocarditis, wounds and abdominal-pelvic
 CC infection) and for screening drugs such as agonists and antagonists. The
 CC nucleic acid is useful for recombinant production of Candida albicans -
 CC derived peptides or antisense polypeptides. Pharmaceutical compositions
 CC and vaccines containing the nucleic acid are useful for preventing or
 CC treating Enterococcus faecium infections. The present sequence represents
 CC one if the disclosed E. faecium proteins.

XX SQ Sequence 177 AA;

Query Match 91.9%; Score 34; DB 7; Length 177;
 Best Local Similarity 71.4%; Pred. No. 9.6e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAPVP 7
 ||:||||
 Db 63 ILAPVP 69

RESULT 34
 ADK34392
 ID ADK34392 standard; protein; 291 AA.

XX AC ADK34392;

XX DT 06-MAY-2004 (first entry)

XX DE Novel human polypeptide SeqID6474.

XX KW antiarthritic; antiparkinsonian; neuroprotective; nootropic;
 KW immunosuppressive; cytostatic; antiporiatic; antiinflammatory;
 KW antibacterial; antiviral; antifungal; antiparasitic; gene therapy;
 KW arthritis; Parkinson's; Alzheimer's; autoimmune disease; cancer;
 KW psoriasis; inflammatory bowel disease; infection; bacteria; virus;
 KW fungus; parasite; human.

XX OS Homo sapiens.

XX PH Key Location/Qualifiers

XX FT Misc-difference 1..291

XX FT /label= OTHER

XX FT /note= "OTHER= All Xaa's in this sequence are unknown
 FT amino acids or the site of a stop codon within the DNA
 FT sequence"

XX PN WO200216439-A2.

XX PD 28-FEB-2002.

XX PF 05-MAR-2001; 2001WO-US004941.

XX PR 07-MAR-2000; 2000US-00519705.

XX PR 19-MAY-2000; 2000US-00574454.

XX PA (HYSE-) HYSEQ INC.

XX PI Tang YT, Liu C, Drmanac RT;

XX DR WPI; 2002-280918/32.

XX PT Isolated polynucleotide encoding bone marrow derived polypeptides useful
 PT for treating, e.g., Parkinson's, Alzheimer's, cancer, arthritis, Crohn's
 PT disease, and inflammatory bowel disease.

XX PS Claim 20; SEQ ID NO 6474; 504pp; English.

XX CC This invention relates to a novel isolated polynucleotide comprising a
 CC nucleotide sequence selected from one of 1680 sequences, a mature protein
 CC coding portion of them, an active domain of them and their complementary
 CC sequences. The invention may be useful for the production of compounds
 CC with an antiarthritic, antiparkinsonian, neuroprotective, nootropic,
 CC immunosuppressive, cytostatic, antiporiatic, antiinflammatory,
 CC antibacterial, antiviral, antifungal or antiparasitic activity. In
 CC addition, the disclosed sequences may be useful for gene therapy. The
 CC polypeptides or their antibodies are useful for treating many diseases
 CC such as arthritis, Parkinson's, Alzheimer's, autoimmune diseases, cancer,
 CC psoriasis, inflammatory bowel disease and infections caused by bacteria,
 CC viruses, fungi or parasites. The present sequence is that of a human
 CC polypeptide of the invention.

XX SQ Sequence 291 AA;

Query Match 91.9%; Score 34; DB 5; Length 291;
 Best Local Similarity 85.7%; Pred. No. 1.5e+03;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAPVP 7

Db 123 ILAPVP 129

RESULT 35
ADV58105
ID ADV58105 standard; peptide; 38 AA.
XX AC ADV58105;
XX DT 10-MAR-2005 (first entry)
XX DE G protein coupled receptor peptide SEQ ID NO 5609.
XX diagnosis; cancer; obesity; diabetes; asthma; inflammation; depression;
KW food; feedstuff; cosmetics; agriculture; animal breeding; GPCR.
XX Unidentified.
XX OS WO2004111636-A2.
XX PN 23-DEC-2004.
XX PD 17-JUN-2004; 2004WO-EP051158.
XX PF 17-JUN-2003; 2003EP-00101775.
XX PR 17-JUN-2003; 2003US-0479061P.
XX PA (VIBV-) VIB VZW.
XX PE (UYGE-) UNIV GENT.
XX PI Kas K, Vandekerckhove J, Krols L;
XX WPI; 2005-057893/06.
XX PT Identifying a peptide combo which corresponds with a family of proteins,
PT useful for diagnosing a variety of diseases, drug development or in
PT agriculture, comprises generating peptides by applying a digest on the
PT family of protein.
XX Example; SEQ ID NO 5609; 265pp; English.
XX The invention relates to a method of identifying a peptide combo which
CC corresponds with a family of proteins where each of the members of the
CC peptide combo is derived from a unique protein from the family. The
CC peptide combo is useful for quantifying specific known splice variants of
CC one or more particular proteins in a sample, for diagnosing complex
CC genetic diseases such as cancer, obesity, diabetes, asthma and
CC inflammation, neuropsychiatric disorders such as depression, for
CC quantifying one to several hundreds of protein disease markers
CC simultaneously leading to a more accurate diagnostic sub-classification,
CC for determining the extent of protein modification in a particular sample
CC of proteins, for tissue-typing analysis, for prenatal testing to detect
CC the presence of a congenital disease or for quantitating protein levels
CC diagnostic of a chromosomal abnormality, for diagnosing immune diseases
CC or neurological diseases, as biomarkers preclinical drug development,
CC development of improved animal models, biomarkers related with
CC toxicology, clinical drug development, guidance marketed drugs,
CC prognostic or diagnostic disease markers, drug target validation and
CC selection, monitoring protein splicing, drug lead profiling, pathway
CC analysis, answering basic disease biology questions, and in the fields of
CC food and feed, cosmetics, agriculture and animal breeding. The present
CC sequence represents a peptide from a G-protein coupled receptor peptide
CC combo.
XX SQ Sequence 38 AA;
Query Match 89.2%; Score 33; DB 9; Length 38;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 LAPPVP 7
| | | | |
Db 22 LAPPVP 27

RESULT 36
ADV59047
ID ADV59047 standard; peptide; 39 AA.
XX AC ADV59047;
XX DT 10-MAR-2005 (first entry)
XX DE G protein coupled receptor peptide SEQ ID NO 6552.
XX diagnosis; cancer; obesity; diabetes; asthma; inflammation; depression;
KW food; feedstuff; cosmetics; agriculture; animal breeding; GPCR.
XX Unidentified.
XX OS WO2004111636-A2.
XX PN 23-DEC-2004.
XX PD 17-JUN-2004; 2004WO-EP051158.
XX PF 17-JUN-2003; 2003EP-00101775.
XX PR 17-JUN-2003; 2003US-0479061P.
XX PA (VIBV-) VIB VZW.
XX PE (UYGE-) UNIV GENT.
XX PI Kas K, Vandekerckhove J, Krols L;
XX WPI; 2005-057893/06.
XX PT Identifying a peptide combo which corresponds with a family of proteins,
PT useful for diagnosing a variety of diseases, drug development or in
PT agriculture, comprises generating peptides by applying a digest on the
PT family of protein.
XX Example; SEQ ID NO 6552; 265pp; English.
XX The invention relates to a method of identifying a peptide combo which
CC corresponds with a family of proteins where each of the members of the
CC peptide combo is derived from a unique protein from the family. The
CC peptide combo is useful for quantifying specific known splice variants of
CC one or more particular proteins in a sample, for diagnosing complex
CC genetic diseases such as cancer, obesity, diabetes, asthma and
CC inflammation, neuropsychiatric disorders such as depression, for
CC quantifying one to several hundreds of protein disease markers
CC simultaneously leading to a more accurate diagnostic sub-classification,
CC for determining the extent of protein modification in a particular sample
CC of proteins, for tissue-typing analysis, for prenatal testing to detect
CC the presence of a congenital disease or for quantitating protein levels
CC diagnostic of a chromosomal abnormality, for diagnosing immune diseases
CC or neurological diseases, as biomarkers preclinical drug development,
CC development of improved animal models, biomarkers related with
CC toxicology, clinical drug development, guidance marketed drugs,
CC prognostic or diagnostic disease markers, drug target validation and
CC selection, monitoring protein splicing, drug lead profiling, pathway
CC analysis, answering basic disease biology questions, and in the fields of
CC food and feed, cosmetics, agriculture and animal breeding. The present
CC sequence represents a peptide from a G-protein coupled receptor peptide
CC combo.
XX SQ Sequence 39 AA;
Query Match 89.2%; Score 33; DB 9; Length 39;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 LAPPVP 7
| | | | |
Db 22 LAPPVP 27

RESULT 37

```

ADC94833
ID ADC94833 standard; protein; 67 AA.
XX
AC ADC94833;
XX
DT 01-JAN-2004 (first entry)
XX
DE E. faecium protein sequence SEQ ID 4460.
XX
XX Vaccine; urinary tract infection; bacteraemia; endocarditis; wound;
KW abdominal-pelvic infection.
XX
KW Enterococcus faecium.
OS
XX
XX US6583275-B1.
PN
XX
XX 24-JUN-2003.
PD
XX
XX 30-JUN-1998; 98US-00107532.
PF
XX 02-JUL-1997; 97US-0051571P.
PR
XX 14-MAY-1998; 98US-0085598P.
PR
XX (GENO-) GENOME THERAPEUTICS CORP.
PA
XX
XX Doucette-Stamm LA, Bush D;
PI
XX
XX WPI; 2003-799836/75.
DR
XX N-PSDB; ADC91179.
DR
XX
XX New isolated nucleic acid derived from Enterococcus faecium encoding an
PT Enterococcus faecium polypeptide useful for detection, prevention and
PT treatment of a pathological condition resulting from a bacterial
PT infection.
XX
XX Example 1; SEQ ID NO 4460; 243pp; English.
PS
XX
XX The invention relates to an isolated nucleic acid derived from
CC Enterococcus faecium encoding an Enterococcus faecium polypeptide having
CC one of 10 fully defined sequences given in the (or comprising 40
CC sequential nucleotides chosen from any of the nucleic acids, its
CC complement or sequences hybridising to it). Also included are a
CC recombinant vector comprising the nucleic acid operably linked to
CC transcription regulatory element, a cell comprising the vector and a
CC single-stranded probe comprising the nucleic acid. The nucleic acids are
CC chosen from 3654 disclosed sequences encoding 3654 disclosed proteins.
CC The nucleic acids is useful for diagnosing pathological conditions
CC resulting from E. faecium bacterial infection (e.g. urinary tract
CC infection, bacteraemia, endocarditis, wounds and abdominal-pelvic
CC infection) and for screening drugs such as agonists and antagonists. The
CC nucleic acid is useful for recombinant production of Candida albicans -
CC derived peptides or antisense polypeptides. Pharmaceutical compositions
CC and vaccines containing the nucleic acid are useful for preventing or
CC treating Enterococcus faecium infections. The present sequence represents
CC one if the disclosed E. faecium proteins.
XX
XX Sequence 67 AA;
SQ
Query Match 89.2%; Score 33; DB 7; Length 67;
Best Local Similarity 85.7%; Pred. No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ILAPPVP 7
DB 39 IPAPPVP 45
| | | | |
| | | | |
RESULT 38
AAY00298
ID AAY00298 standard; protein; 93 AA.
XX
XX
AC AAY00298;
XX
DT 04-MAY-1999 (first entry)
XX
DE Human secreted protein encoded by gene 41.
XX
XX Human; secreted protein; fusion protein; gene therapy; protein therapy;
KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
KW developmental abnormality; foetal deficiency; blood; allergy; renal;
KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
XX
XX Homo sapiens.
OS
XX
XX WO9906423-A1.
PN
XX
XX 11-FEB-1999.
PD
XX
XX 29-JUL-1998; 98WO-US015949.
PF
XX 30-JUL-1997; 97US-0054209P.
PR
XX 30-JUL-1997; 97US-0054211P.
PR
XX 30-JUL-1997; 97US-0054212P.
PR
XX 30-JUL-1997; 97US-0054213P.
PR
XX 30-JUL-1997; 97US-0054214P.
PR
XX 30-JUL-1997; 97US-0054215P.
PR
XX 30-JUL-1997; 97US-0054217P.
PR
XX 30-JUL-1997; 97US-0054218P.
PR
XX 30-JUL-1997; 97US-0054234P.
PR
XX 30-JUL-1997; 97US-0054236P.
PR
XX 18-AUG-1997; 97US-0055968P.
PR
XX 18-AUG-1997; 97US-0055969P.
PR
XX 18-AUG-1997; 97US-0055972P.
PR
XX 19-AUG-1997; 97US-0056534P.
PR
XX 19-AUG-1997; 97US-0056543P.
PR
XX 19-AUG-1997; 97US-0056554P.
PR
XX 19-AUG-1997; 97US-0056561P.
PR
XX 19-AUG-1997; 97US-0056727P.
PR
XX 19-AUG-1997; 97US-0056729P.
PR
XX 19-AUG-1997; 97US-0056730P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX
XX Ruben SM, Feng P, Lafleur DM, Moore PA, Shi Y, Kyaw H, Li Y;
PI Zeng Z, Carter KC, Endress GA, Wei Y, Fan P, Rosen CA;
XX
XX WPI; 1999-153691/13.
DR N-PSDB; AAX20452.
DR
XX
XX New isolated human genes and the secreted polypeptides they encode -
PT useful for diagnosis and treatment of e.g. cancers, neurological
PT disorders, immune diseases, inflammation or blood disorders.
XX
XX Claim 11; Page 273; 312pp; English.
PS
XX
XX This sequence represents a secreted human protein encoded by the gene
CC clone detailed in the descriptor line. The gene can be used to generate
CC fusion proteins by linking to the gene to a human immunoglobulin Fc
CC portion (e.g. AAX20403) for increasing the stability of the fused protein
CC as compared to the human protein only. The invention relates to 83 novel
CC genes and their fragments (nucleic acid sequences: AAX20412-X20499; amino
CC acid sequences AAY00258-Y00377) which are useful for preventing, treating
CC or ameliorating medical conditions e.g. by protein or gene therapy. Also,
CC pathological conditions can be diagnosed by determining the amount of the
CC new polypeptides in a sample or by determining the presence of mutations
CC in the new polynucleotides. Specific uses are described for each of the
CC 83 polynucleotides, based on which tissues they are most highly expressed
CC in (see AAX20412 for described uses)
XX
XX Sequence 93 AA;
SQ
Query Match 89.2%; Score 33; DB 2; Length 93;

```

PT	preventing, treating or ameliorating a medical condition e.g., cancer.
XX	Claim 11; SEQ ID NO 142; 213pp; English.
XX	The invention describes novel isolated human nucleic acids. The nucleic acid is useful for preparing a medicament for preventing, treating or ameliorating a medical condition e.g., cancer, and in gene therapy. This is the amino acid sequence of a novel human secreted protein of the invention.
XX	Sequence 93 AA;
SQ	Query Match 89.2%; Score 33; DB 7; Length 93; Best Local Similarity 100.0%; Pred. No. 7.5e+02; Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy	2 LAPPVP 7
Db	8 LAPPVP 13
RESULT 40	
ADG90146	
ID	ADG90146 standard; protein; 93 AA.
XX	AC ADG90146;
XX	11-MAR-2004 (first entry)
DT	Human protein from secreted protein gene 41.
XX	Secreted protein; gene therapy; neural disorder; immune system disorders; muscular disorder; reproductive disorder; gastrointestinal disorder; pulmonary disorder; cardiovascular disorder; renal disorder; proliferative disorder; cancer; systemic lupus erythematosus; rheumatoid arthritis; multiple sclerosis; thyroiditis; anaemia; Grave's disease; diabetes; hepatitis; asthma; allergy; nephritis; Parkinson's disease; Alzheimer's disease; atherosclerosis; myocardial infarction; AIDS; infection; human.
OS	Homo sapiens.
XX	US2003166541-A1.
PN	04-SEP-2003.
PD	04-JUN-2002; 2002US-00160162.
PP	30-JUL-1997; 97US-0054209P.
PR	30-JUL-1997; 97US-0054211P.
PR	30-JUL-1997; 97US-0054212P.
PR	30-JUL-1997; 97US-0054213P.
PR	30-JUL-1997; 97US-0054214P.
PR	30-JUL-1997; 97US-0054215P.
PR	30-JUL-1997; 97US-0054216P.
PR	30-JUL-1997; 97US-0054217P.
PR	30-JUL-1997; 97US-0054218P.
PR	30-JUL-1997; 97US-0054234P.
PR	30-JUL-1997; 97US-0054236P.
PR	18-AUG-1997; 97US-0055968P.
PR	18-AUG-1997; 97US-0055969P.
PR	18-AUG-1997; 97US-0055972P.
PR	19-AUG-1997; 97US-0056534P.
PR	19-AUG-1997; 97US-0056543P.
PR	19-AUG-1997; 97US-0056554P.
PR	19-AUG-1997; 97US-0056561P.
PR	19-AUG-1997; 97US-0056727P.
PR	19-AUG-1997; 97US-0056729P.
PR	19-AUG-1997; 97US-0056730P.
PR	29-JUL-1998; 98WO-US015949.
PR	26-JAN-1999; 99US-00236557.
PR	05-JUN-2001; 2001US-0295558P.
XX	(HUMA-) HUMAN GENOME SCI INC.
PA	
PT	New nucleic acid molecule, useful for preparing a medicament for

XX Ruben SM, Feng P, Lafleur DW, Moore PA, Shi Y, Kyaw H, Li Y;
 PI Zeng Z, Carter KC, Endress GA, Wei Y, Fan P, Rosen CA;
 XX WPI; 2003-874923/81.
 DR N-PSDB; ADG90055.
 XX
 PT Nucleic acids encoding 83 secreted polypeptides, useful for preventing,
 PT diagnosing and treating disorders related to their aberrant expression
 PT and activity.
 XX
 PS Claim 11; SEQ ID NO 142; 308pp; English.
 XX
 CC The invention relates to an isolated nucleic acid molecule encoding a
 CC secreted protein that is at least 95% identical to a polynucleotide
 CC fragment of any of the nucleotide sequences listed in table 1A of the
 CC specification, which is hybridizable to the nucleotide sequences, a
 CC polynucleotide encoding a polypeptide (or a polypeptide fragment, domain
 CC or epitope of any of the amino acid sequences) listed in table 1A of the
 CC specification, a polynucleotide which is an (allelic) variant of the
 CC nucleotide sequences listed in the specification, a polynucleotide which
 CC encodes a species homologue of the above amino acid sequences, a
 CC polynucleotide capable of hybridizing under stringent conditions to any
 CC of the above polynucleotides, where the polynucleotide does not hybridize
 CC under stringent conditions to a nucleic acid molecule having a nucleotide
 CC sequence of only A or T residues. Also included are a recombinant vector
 CC comprising the above nucleic acid molecule, making a recombinant host
 CC cell comprising the above nucleic acid molecule, an isolated polypeptide
 CC comprising a sequence that is at least 95% identical to the polypeptide
 CC (or its fragment, domain, epitope, secreted form, (allelic) variant or
 CC homologue) encoded by the above nucleic acid molecule, an isolated
 CC antibody that binds specifically to the above polypeptide, a recombinant
 CC host cell produced by the above method and that expresses the above
 CC polypeptide, making an isolated polypeptide, preventing, treating or
 CC ameliorating a medical condition, diagnosing a pathological condition or
 CC a susceptibility to a pathological condition in a subject, identifying a
 CC binding partner to the above polypeptide, the gene corresponding to the
 CC cDNA sequence given in the specification, and identifying an activity in
 CC a biological assay. The nucleic acid molecule and polypeptide are useful
 CC in diagnosing, preventing, prognosing or treating diseases or disorders
 CC associated with aberrant expression and/or activity of the above
 CC polypeptide, such as neural disorders, immune system disorders, muscular
 CC disorders, reproductive disorders, gastrointestinal disorders, pulmonary
 CC disorders and/or cancers. In particular, these diseases are systemic
 CC lupus erythematosus, rheumatoid arthritis, multiple sclerosis,
 CC thyroiditis, anaemia, Grave's disease, diabetes, hepatitis, asthma,
 CC allergies, nephritis, Parkinson's disease, Alzheimer's disease,
 CC atherosclerosis, myocardial infarction, AIDS and infections. The methods
 CC may be used for identifying agonists and antagonists of the
 CC polynucleotide and polypeptide. The present sequence is a protein from
 CC one of the 83 disclosed secreted protein genes.
 XX
 SX Sequence 93 AA;
 SQ
 Query Match 89.2%; Score 33; DB 7; Length 93;
 Best Local Similarity 100.0%; Pred. No. 7.5e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 LAPPVP 7
 DB 8 LAPPVP 13
 |||||
 RESULT 41
 ID ADY25486
 AC ADY25486 standard; protein; 93 AA.
 XX
 AC ADY25486;
 XX
 DT 05-MAY-2005 (first entry)
 XX
 DE Novel human secreted protein SeqID142.

XX cancer; cytostatic; arthritis; antiarthritic; asthma; antiasthmatic;
 KW acquired immune deficiency syndrome; rheumatoid arthritis; antirheumatic;
 KW inflammatory bowel disease; antiinflammatory; gastrointestinal-gen.;
 KW sepsis; antibacterial; immunosuppressive; acne; antiiseborrheic;
 KW dermatological; psoriasis; antipsoriatic; atherosclerosis;
 KW antiarteriosclerotic; cerebrovascular ischemia; cerebroprotective;
 KW vasotrophic; thrombosis; wound healing; vulnery; Alzheimer's disease;
 KW neuroprotective; nootropic; parkinson's disease; antiparkinsonian; autism;
 KW obsessive-compulsive disorder; tranquilizer; graft versus host disease;
 KW immune disorder; hematological disease; inflammation; infection;
 KW hyperproliferative disorders; renal disease; nephrotropic;
 KW cardiovascular disease; cardiovascular-gen.; respiratory disorder;
 KW neurological disease; neuroprotective; endocrine disease;
 KW reproductive disorders (general); gynecological.
 XX
 XX Homo sapiens.
 OS
 XX US2005037467-A1.
 PN
 PD 17-FEB-2005.
 XX
 XX 09-SEP-2004; 2004US-00936773.
 PF
 XX 30-JUL-1997; 97US-0054209P.
 XX 30-JUL-1997; 97US-0054211P.
 PR 30-JUL-1997; 97US-0054212P.
 PR 30-JUL-1997; 97US-0054213P.
 PR 30-JUL-1997; 97US-0054214P.
 PR 30-JUL-1997; 97US-0054215P.
 PR 30-JUL-1997; 97US-0054217P.
 PR 30-JUL-1997; 97US-0054218P.
 PR 30-JUL-1997; 97US-0054234P.
 PR 30-JUL-1997; 97US-0054236P.
 PR 18-AUG-1997; 97US-0055968P.
 PR 18-AUG-1997; 97US-0055969P.
 PR 18-AUG-1997; 97US-0055972P.
 PR 19-AUG-1997; 97US-0056534P.
 PR 19-AUG-1997; 97US-0056543P.
 PR 19-AUG-1997; 97US-0056554P.
 PR 19-AUG-1997; 97US-0056561P.
 PR 19-AUG-1997; 97US-0056727P.
 PR 19-AUG-1997; 97US-0056729P.
 PR 19-AUG-1997; 97US-0056730P.
 PR 29-JUL-1998; 98WO-US015949.
 PR 26-JAN-1999; 99US-00236557.
 PR 21-SEP-2000; 2000US-00666984.
 PR 30-MAR-2001; 2001US-00820649.
 PR 05-JUN-2001; 2001US-0295558P.
 PR 04-JUN-2002; 2002US-00160162.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA
 XX Ruben SM, Feng P, Lafleur DW, Moore PA, Shi Y, Kyaw H, Li Y;
 PI Zeng Z, Carter KC, Endress GA, Wei Y, Fan P, Rosen CA;
 XX WPI; 2005-161941/17.
 DR N-PSDB; ADY25395.
 XX
 PT New nucleic acid molecule encoding human secreted protein, useful for
 PT preventing, treating, or ameliorating immune system, blood, inflammatory,
 PT infectious, cardiovascular, respiratory, neurological, endocrine, or
 PT reproductive disorders.
 XX
 PS Claim 11; SEQ ID NO 142; 309pp; English.
 XX
 CC The invention comprises the amino acid and coding sequences of novel
 CC human secreted proteins and peptides. The DNA and protein sequences of
 CC the invention are useful for treating or ameliorating: cancer, arthritis,
 CC asthma, AIDS, rheumatoid arthritis, inflammatory bowel disease, sepsis,
 CC acne, psoriasis, atherosclerosis, stroke, thrombosis, wound healing,
 CC Alzheimer's disease, Parkinson's disease, autism, obsessive compulsive
 CC disorder, graft-versus-host diseases, immune system disorders, blood

CC disorders, inflammatory conditions, infectious diseases,
CC hyperproliferative disorders, renal disorders, cardiovascular disorders,
CC respiratory disorders, neurological disorders, endocrine disorders, or
CC reproductive disorders. The present sequence is that of a novel human
CC secreted protein of the invention.
XX
XX

Sequence 93 AA;

Query Match 89.2%; Score 33; DB 9; Length 93;
Best Local Similarity 100.0%; Pred. No. 7.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LAPPVP 7
|||||
DB 8 LAPPVP 13

RESULT 42
ADP60277
ID ADF60277 standard; protein; 95 AA.
XX
AC ADF60277;
DT 12-FEB-2004 (first entry)
XX
XX Human contig polypeptide sequence SEQ ID NO:2644.
XX
XX biological activity; genetic engineering; hybridisation probe; oligomer;
KW primer; chromosome mapping; gene mapping; recombinant protein production;
KW human.
XX
XX Homo sapiens.
XX
XX WO2003080795-A2.
XX
XX 02-OCT-2003.
XX
XX 09-AUG-2002; 2002WO-US025485.
XX
XX 09-AUG-2001; 2001US-0311261P.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Yang Y, Wang Z, Weng G, Ma Y;
XX
XX WPI; 2003-876918/81.
XX
XX N-PSDB; ADF59825.

New polynucleotides, useful as hybridization probes, oligomers or
primers, for chromosome or gene mapping, for the recombinant production
of proteins, and for generating antisense DNA or RNA.
XX
XX Example 3; SEQ ID NO 2644; 571bp; English.
XX
XX The present invention describes isolated polynucleotide sequences (I),
XX which encode polypeptides (II) with biological activity. Also described:
CC (1) a vector comprising (i); (2) an expression vector comprising (i); (3)
CC a host cell genetically engineered to comprise (i) which is operatively
CC associated with a regulatory sequence that modulates expression of (i) in
CC the host cell; (4) a polypeptide (II) encoded by (i); (5) a composition
CC comprising the polypeptide of (4) and a carrier; (6) an antibody directed
CC against the polypeptide of (4); (7) detecting (i) or the polypeptide of
CC (4) in a sample; (8) identifying a compound that binds to the polypeptide
CC of (4); (9) producing the polypeptide of (4); and (10) a collection of
CC polynucleotides comprising at least one of the polynucleotide sequences
CC (I). The polynucleotides (I) can be used as hybridisation probes,
CC oligomers or primers, for chromosome or gene mapping, for the recombinant
CC production of proteins, and for generating antisense DNA or RNA. The
CC present sequence represents a human contig polypeptide sequence, which is
CC used in an example from the present invention.

Sequence 95 AA;

Query Match 89.2%; Score 33; DB 7; Length 95;
Best Local Similarity 85.7%; Pred. No. 7.7e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ILAPPVP 7
|||||
DB 53 ILAPPVP 59

RESULT 43
ABBI7509
ID ABB17509 standard; protein; 112 AA.

XX
AC ABB17509;
XX
XX 23-JAN-2002 (first entry)

XX Human nervous system related polypeptide SEQ ID NO 6166.
XX
XX Human; neurotropic; neuroprotective; cytostatic; dermatological; virucide;
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnery;
KW antiparkinsonian; antisickling; antianaemic; antiarthritic; cancer;
KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
KW antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;
KW antiparasitic; cardiant; immune disorder; cardiovascular disorder;
KW neurological disease; infection; nephrotropic; gene therapy; vaccine.

XX Homo sapiens.

XX WO200159063-A2.

XX 16-AUG-2001.

XX 17-JAN-2001; 2001WO-US001334.

XX 31-JAN-2000; 2000US-0179065P.

XX 04-FEB-2000; 2000US-0180628P.

XX 24-FEB-2000; 2000US-0184664P.

XX 02-MAR-2000; 2000US-0186350P.

XX 16-MAR-2000; 2000US-0189874P.

XX 17-MAR-2000; 2000US-0190076P.

XX 18-APR-2000; 2000US-0198123P.

XX 19-MAY-2000; 2000US-0205515P.

XX 07-JUN-2000; 2000US-0209467P.

XX 28-JUN-2000; 2000US-0214886P.

XX 30-JUN-2000; 2000US-0215135P.

XX 07-JUL-2000; 2000US-0216647P.

XX 07-JUL-2000; 2000US-0216880P.

XX 11-JUL-2000; 2000US-0217487P.

XX 11-JUL-2000; 2000US-0217496P.

XX 14-JUL-2000; 2000US-0218290P.

XX 26-JUL-2000; 2000US-0220963P.

XX 26-JUL-2000; 2000US-0220964P.

XX 14-AUG-2000; 2000US-0224518P.

XX 14-AUG-2000; 2000US-0224519P.

XX 14-AUG-2000; 2000US-0225213P.

XX 14-AUG-2000; 2000US-0225214P.

XX 14-AUG-2000; 2000US-0225266P.

XX 14-AUG-2000; 2000US-0225267P.

XX 14-AUG-2000; 2000US-0225268P.

XX 14-AUG-2000; 2000US-0225270P.

XX 14-AUG-2000; 2000US-0225447P.

XX 14-AUG-2000; 2000US-0225757P.

XX 14-AUG-2000; 2000US-0225758P.

XX 14-AUG-2000; 2000US-0225759P.

XX 18-AUG-2000; 2000US-0226279P.

XX 22-AUG-2000; 2000US-0226681P.

XX 22-AUG-2000; 2000US-0226688P.

XX 23-AUG-2000; 2000US-0227182P.

XX 23-AUG-2000; 2000US-0227009P.

XX 30-AUG-2000; 2000US-0228924P.

XX 01-SEP-2000; 2000US-0228287P.

XX 01-SEP-2000; 2000US-0229343P.

XX 29-JUL-2004 (first entry)
 XX DT
 XX DE Pseudomonas aeruginosa polypeptide #628.
 XX KW
 XX KW Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
 XX OS
 XX OS Pseudomonas aeruginosa.
 XX XX
 XX PN US6551795-B1.
 XX XX
 XX PD 22-APR-2003.
 XX XX
 XX PF 18-FEB-1999; 99US-00252991.
 XX XX
 XX PR 18-FEB-1998; 98US-0074788P.
 XX PR 27-JUL-1998; 98US-0094190P.
 XX XX
 XX PA (GENO-) GENOME THERAPEUTICS CORP.
 XX XX
 XX PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;
 XX DR WPI; 2003-615309/58.
 XX DR N-PSDB; ABD02024.
 XX XX
 XX PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide, of
 XX PT useful as molecular targets for diagnostics, prophylaxis and treatment of
 XX PT pathological conditions resulting from bacterial infection.
 XX PS
 XX PS Disclosure; SEQ ID NO 17199; 455pp; English.
 XX CC
 CC The invention relates to Pseudomonas aeruginosa polypeptides and the
 CC polynucleotides encoding them. The sequences are useful in diagnosis and
 CC therapy of pathological conditions, as molecular targets for diagnostics,
 CC prophylaxis and treatment of pathological conditions resulting from a
 CC bacterial infection, for evaluating a compound, such as a polypeptide,
 CC for the ability to bind a P. aeruginosa nucleic acid, as components of
 CC effective antibacterial targets, as targets for antibacterial drugs,
 CC including anti-P. aeruginosa drugs, as templates for recombinant
 CC production of P. aeruginosa-derived peptides or polypeptides, as target
 CC components for diagnosis and/or treatment of P. aeruginosa-caused
 CC infection, and in detection of P. aeruginosa sequences or other sequences
 CC of Pseudomonas species using biochip technology. Sequences ABO67826-
 CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The
 CC sequence data for this patent did not form part of the printed
 CC specification but was obtained in electronic format from USPTO at
 CC seqdata.uspto.gov/sequence.html
 XX SQ Sequence 161 AA;
 Query Match 89.2%; Score 33; DB 7; Length 161;
 Best Local Similarity 100.0%; Pred. No. 1.3e+03;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 LAPPVP 7
 DB 23 LAPPVP 28
 RESULT 45
 ABG27197
 ID ABG27197 standard; protein; 190 AA.
 XX AC
 XX AC ABG27197;
 XX DT
 XX DT 18-FEB-2002 (first entry)
 DE Novel human diagnostic protein #27188.
 XX KW
 XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX KW
 XX KW Homo sapiens.

XX WO200175067-A2.
 XX PN
 XX PD 11-OCT-2001.
 XX XX
 XX PF 30-MAR-2001; 2001WO-US008631.
 XX XX
 XX PR 31-MAR-2000; 2000US-00540217.
 XX PR 23-AUG-2000; 2000US-00649167.
 XX XX
 XX PA (HYSE-) HYSEQ INC.
 XX XX
 XX PI Drmanac RT, Liu C, Tang YT;
 XX XX
 XX DR WPI; 2001-639362/73.
 XX DR N-PSDB; AAS91384.
 XX XX
 XX PT New isolated polynucleotide and encoded polypeptides, useful in
 XX PT diagnostics, forensics, gene mapping, identification of mutations
 XX PT responsible for genetic disorders or other traits and to assess
 XX PT biodiversity.
 XX PS
 XX PS Claim 20; SEQ ID NO 57556; 103pp; English.
 XX CC
 CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
 CC amino acid sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX SQ Sequence 190 AA;
 Query Match 89.2%; Score 33; DB 4; Length 190;
 Best Local Similarity 71.4%; Pred. No. 1.5e+03;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ILAPPVP 7
 DB 32 VLAPPLP 38
 RESULT 46
 ADX74877
 ID ADX74877 standard; protein; 191 AA.
 XX AC
 XX AC ADX74877;
 XX DT
 XX DT 21-APR-2005 (first entry)
 XX XX
 XX DE Plant full length insert polypeptide seqid 44243.
 XX KW
 KW plant protectant; plant growth regulant; gene therapy; plant;
 KW recombinant DNA construct; physical array; plant breeding marker;
 KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
 KW extreme osmotic condition; pathogen tolerance; pest tolerance;
 KW growth rate; cell cycle pathway; disease resistance;
 KW galactomannan production; lignin production; plant growth regulator;

KW yield; plant growth; plant development; seed oil; protein yield;
 KW protein content.

XX Unidentified.

XX US2004034888-A1.

XX 19-FEB-2004.

XX 28-APR-2003; 2003US-00425114.

XX 06-MAY-1999; 99US-00304517.

XX 05-NOV-2001; 2001US-00985678.

XX (LIU/J) LIU J.

XX (ZHOU/J) ZHOU Y.

XX (KOVA/) KOVALIC D K.

XX (SCRE/) SCREEN S E.

XX (TABA/) TABASKA J E.

XX (CAO/Y) CAO Y.

XX Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;

XX WPI; 2004-180133/17.

XX New recombinant DNA construct, useful for improving plant tolerance to
 PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
 PT pests, for conferring increased resistance to plant disease, or for
 PT improving yield.

XX Claim 1; SEQ ID NO 44243; 15pp; English.

XX The invention describes a recombinant DNA construct comprising a
 CC polynucleotide consisting of a sequence encoding an amino acid sequence
 CC available in electronic form from the US patent office at
 CC ftp.seidata.uspto.gov/sequence.html?DocID:2004034888. The polynucleotide
 CC of the invention are also useful in physical arrays of molecules and as
 CC plant breeding markers. The recombinant DNA construct is useful for
 CC improving plant tolerance to cold, heat, drought, herbicides, extreme
 CC osmotic conditions, pathogens or pests, for manipulating growth rate in
 CC plant cells by modification of the cell cycle pathway, for conferring
 CC increased resistance to plant disease, for producing galactomannan,
 CC lignin or plant growth regulators, for increasing the rate of homologous
 CC recombination in plants, for improving yield by modification of
 CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
 CC or by providing improved plant growth and development under at least one
 CC stress condition or for modifying seed oil or protein yield and/or
 CC content. This is the amino acid sequence of a plant full length insert
 CC polypeptide that can be used in the recombinant DNA construct of the
 CC invention.

XX Sequence 191 AA;

Query Match 89.2%; Score 33; DB 8; Length 191;
 Best Local Similarity 100.0%; Pred. No. 1.5e+03; Indels 0; Gaps 0;
 Matches 6; Conservative 0; Mismatches 0;

QY 2 LAPPVP 7

Db 105 LAPPVP 110

RESULT 47

ADC31836

ID ADC31836 standard; protein; 198 AA.

XX ADC31836;

XX 18-DEC-2003 (first entry)

XX Human novel polypeptide sequence, SEQ ID NO:1918.

XX Human; diagnostic; drug screening; forensics; gene mapping;

KW biodiversity assessment; Parkinson's disease; Alzheimer's disease;
 KW neurodegenerative diseases; anaemia; platelet disorder; wound; burns;
 KW ulcers; osteoporosis; autoimmune disease; cancer;
 KW molecular weight marker; food supplement; antiparkinsonian; nootropic;
 KW neuroprotective; antianaemic; anticoagulant; thrombolytic; vulnary;
 KW antiulcer; osteopathic; immunosuppressive; antiinflammatory; cytostatic;
 KW gene therapy.

XX Homo sapiens.

XX WO2003029271-A2.

XX 10-APR-2003.

XX 24-SEP-2002; 2002WO-US030474.

XX 24-SEP-2001; 2001US-0324631P.

XX (HYSE-) HYSEQ INC.

XX Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;

XX Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Wang G;

XX Haley-Vicente D, Drmanac RT;

XX WPI; 2003-371981/35.

XX N-PSDB; ADC30865.

XX New polynucleotide and polypeptide useful for diagnosing, preventing or
 PT treating conditions such as neurodegenerative diseases, anemias, platelet
 PT disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
 PT cancer.

XX Claim 20; SEQ ID NO 1918; 1185pp; English.

XX The invention relates to 971 novel human cDNA sequences (ADC29919-
 CC ADC30899) and the polypeptides they encode (ADC30890-ADC31860). The
 CC invention also relates to nucleic acid sequences over 99% identical with
 CC the novel human cDNAs. The invention additionally encompasses expression
 CC vectors and host cells comprising a nucleic acid of the invention; the
 CC recombinant production of a polypeptide of the invention; an antibody
 CC against a polypeptide of the invention; a method of detecting
 CC polynucleotides or polypeptides of the invention; and methods of
 CC identifying a compound which binds to a polypeptide of the invention. The
 CC invention further discloses methods of preventing, treating or
 CC ameliorating a medical condition; kits comprising polynucleotide probes
 CC and/or monoclonal antibodies for carrying out the methods of the
 CC invention; methods for the identification of compounds that modulate the
 CC expression or activity of the polynucleotide and/or polypeptide; and 767
 CC contig sequences corresponding to the cDNA sequences of the invention
 CC (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628
 CC -ADC33394). The nucleic acids and polypeptides of the invention are
 CC useful in diagnostics, drug screening, forensics, gene mapping, in the
 CC identification of mutations responsible for genetic disorders or other
 CC traits, for assessing biodiversity, and in producing many other types of
 CC data and products dependent on DNA and amino acid sequences. They are
 CC also used for treating diseases such as Parkinson's disease, Alzheimer's
 CC disease and other neurodegenerative diseases, anaemia, platelet
 CC disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
 CC cancer. The nucleic acids may also be used as hybridisation probes or
 CC primers, and in the recombinant production of a protein. The polypeptides
 CC are also useful in generating antibodies, as molecular weight markers,
 CC and as food supplements. The present sequence represents a specifically
 CC claimed human polypeptide sequence of the invention. Note: The sequence
 CC data for this patent did not form part of the printed specification, but
 CC was obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 198 AA;

Query Match 89.2%; Score 33; DB 7; Length 198;
 Best Local Similarity 85.7%; Pred. No. 1.5e+03;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;


```

OY 1 ILAPPVP 7
Db 48 ILGPPVP 54

RESULT 48
ADD26860
ID ADD26860 standard; protein; 207 AA.
XX
AC ADD26860;
XX
DT 15-JAN-2004 (first entry)
XX
DE Human adipocyte Selected Interacting Domain (SID) prey protein #243.
XX
KW Adipocyte; protein-protein interaction; protein complex;
KW bait-prey complex; Selected Interacting Domain; SID; drug screening;
KW drug discovery; metabolic disease; obesity; lipodystrophy;
KW diabetes mellitus; type 2; non-insulin dependent; NIDDM;
KW adipogenesis modulation; gene therapy; human.
XX
OS Homo sapiens.
XX
PN WO200290544-A2.
XX
PD 14-NOV-2002.
XX
PF 03-MAY-2002; 2002WO-EP006333.
XX
PR 04-MAY-2001; 2001US-0288885P.
XX
PA (HYBR-) HYBRIGENICS.
XX
PI (LYNX-) LYNX THERAPEUTICS INC.
XX
PI Legrain P, Whiteside S, Mao J, Khrebtukova I, Luo S;
XX
WPI; 2003-111975/10.
DR N-PSDB; ADD26858.
XX
XX
PT New complex of bait and prey between two polypeptides or polynucleotides
PT encoding the two polypeptides of adipocytes, useful for selecting a
PT modulating compound that inhibits or activates protein-protein
PT interactions.
XX
PS Claim 6; SEQ ID NO 243; 232pp; English.
XX
CC The invention relates to a bait-prey complex between two adipocyte
CC polypeptides, or between two polynucleotides encoding adipocyte
CC polypeptides. The invention also relates to selected interacting Domain
CC (SID) polypeptides which interact with selected bait polypeptides;
CC polynucleotides encoding SID polypeptides; vectors comprising SID
CC polynucleotides; recombinant host cells comprising an adipocyte
CC polynucleotide or a SID-encoding vector; a method of selecting for a
CC compound which modulates interactions between adipocyte polypeptides;
CC adipocyte modulator compounds identified using the method; a
CC pharmaceutical composition comprising an adipocyte modulator, or a SID-
CC encoding vector or host cell; and a protein chip comprising adipocyte
CC bait polypeptides. The bait-prey complexes of the invention are useful
CC for selecting a compound that inhibits or activates protein-protein
CC interactions between adipocyte polypeptides. The modulatory compounds
CC identified can be used in the treatment of metabolic diseases such as
CC obesity, lipodystrophy and type 2 diabetes mellitus, and in the
CC modulation of adipogenesis. The present sequence represents a
CC specifically claimed adipocyte SID prey polypeptide of the invention.
XX
SQ Sequence 207 AA;

Query Match 89.2%; Score 33; DB 7; Length 207;
Best Local Similarity 100.0%; Pred. No. 1.6e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 LAPPVP 7
Db 48 ILGPPVP 54

RESULT 49
AAU14624
ID AAU14624 standard; protein; 267 AA.
XX
AC AAU14624;
XX
DT 24-OCT-2001 (first entry)
XX
DE Novel bone marrow polypeptide #23.
XX
KW Bone marrow; diagnostic; therapeutic; gene therapy; antigenic;
KW haematopoiesis; myeloid; lymph cell disorder; tissue regeneration;
KW wound healing; nutritional supplement; immune disorder;
KW severe combined immunodeficiency; SCID.
XX
OS Homo sapiens.
XX
PN WO200157187-A2.
XX
PD 09-AUG-2001.
XX
PF 05-FEB-2001; 2001WO-US003782.
XX
PR 03-FEB-2000; 2000US-00496914.
PR 20-JUN-2000; 2000US-00598075.
PR 19-JUL-2000; 2000US-00620325.
PR 30-NOV-2000; 2000US-0250683P.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Ford JE, Boyle BJ, Tang YT, Liu C, Asundi V, Zhou P, Xue AJ;
PI Ren F, Drmanac RT;
XX
WPI; 2001-488875/53.
DR N-PSDB; AAS22929.
XX
XX
PT Nucleic acids encoding bone marrow polypeptides, useful in diagnostic and
PT gene therapy.
XX
PS Claim 10; Page 222; 392pp; English.
XX
CC AAU14602-AAU14794 represent novel bone marrow polypeptides of the
CC invention. The proteins and corresponding coding sequences may be used in
CC the prevention, diagnosis and treatment of diseases associated with
CC inappropriate bone marrow polypeptide expression. For example, to treat
CC disorders associated with decreased expression by rectifying mutations or
CC deletions in a patient's genome that affect the activity of the
CC polypeptides by expressing inactive proteins or to supplement the
CC patient's own production of the polypeptide. Additionally, the nucleic
CC acids may be used to produce the polypeptides, by inserting the nucleic
CC acids into a host cell and culturing the cell to express the protein. The
CC nucleic acid and its complementary sequences may also be used as DNA
CC probes in diagnostic assays to detect and quantitate the presence of
CC similar nucleic acid sequences in samples, and therefore which patients
CC may be in need of restorative therapy. The proteins may also be used as
CC antigens in the production of antibodies against bone marrow proteins and
CC in assays to identify modulators of their expression and activity. The
CC anti-bone marrow protein antibodies and antagonists may also be used to
CC down regulate expression and activity. The antibodies may also be used as
CC diagnostic agents for detecting the presence of the protein in samples
CC (e.g. by enzyme linked immunosorbant assay (ELISA)). The proteins may be
CC used to regulate haematopoiesis activity, and consequently in the
CC treatment of myeloid or lymph cell disorders; in tissue regeneration,
CC such as wound healing; as a nutritional supplement; and in treatment of
CC immune disorders such as severe combined immunodeficiency (SCID)
XX
SQ Sequence 267 AA;

Query Match 89.2%; Score 33; DB 4; Length 267;
Best Local Similarity 100.0%; Pred. No. 2e+03;

```

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LAPPVP 7
 Db 133 LAPPVP 138
 |||||

RESULT 50
 ADX77177
 ID ADX77177 standard; protein; 285 AA.
 XX
 AC ADX77177;
 XX
 DT 21-APR-2005 (first entry)
 XX
 DE Plant full length insert polypeptide seqid 46543.
 XX
 KW plant protectant; plant growth regulant; gene therapy; plant;
 KW recombinant DNA construct; physical array; plant breeding marker;
 KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
 KW extreme osmotic condition; pathogen tolerance; pest tolerance;
 KW growth rate; cell cycle pathway; disease resistance;
 KW galactomannan production; lignin production; plant growth regulator;
 KW yield; plant growth; plant development; seed oil; protein yield;
 KW protein content.
 XX
 OS Unidentified.
 XX
 PN US2004034888-A1.
 XX
 PD 19-FEB-2004.
 XX
 PF 28-APR-2003; 2003US-00425114.
 XX
 PR 06-MAY-1999; 99US-00304517.
 PR 05-NOV-2001; 2001US-00985678.
 XX
 PA (LIU/) LIU J.
 PA (ZHOU/) ZHOU Y.
 PA (KOVA/) KOVALIC D K.
 PA (SCRE/) SCREEN S E.
 PA (TABA/) TABASKA J E.
 PA (CAOY/) CAO Y.
 XX
 PI Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
 XX
 DR WPI; 2004-180133/17.
 XX
 PT New recombinant DNA construct, useful for improving plant tolerance to
 PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
 PT pests, for conferring increased resistance to plant disease, or for
 PT improving yield.
 XX
 PS Claim 1; SEQ ID NO 46543; 15pp; English.
 XX
 CC The invention describes a recombinant DNA construct comprising a
 CC polynucleotide consisting of a sequence encoding an amino acid sequence
 CC available in electronic form from the US patent office at
 CC ftp.segdata.uspto.gov/sequence.html?DocID:2004034888. The polynucleotide
 CC of the invention are also useful in physical arrays of molecules and as
 CC plant breeding markers. The recombinant DNA construct is useful for
 CC improving plant tolerance to cold, heat, drought, herbicides, extreme
 CC osmotic conditions, pathogens or pests, for manipulating growth rate in
 CC plant cells by modification of the cell cycle pathway, for conferring
 CC increased resistance to plant disease, for producing galactomannan,
 CC lignin or plant growth regulators, for increasing the rate of homologous
 CC recombination in plants, for improving yield by modification of
 CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
 CC or by providing improved plant growth and development under at least one
 CC stress condition or for modifying seed oil or protein yield and/or
 CC content. This is the amino acid sequence of a plant full length insert
 CC polypeptide that can be used in the recombinant DNA construct of the
 CC invention.

XX Sequence 285 AA;
 SQ
 Query Match 89.2%; Score 33; DB 8; Length 285;
 Best Local Similarity 100.0%; Pred. No. 2.1e+03;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 LAPPVP 7
 |||||
 Db 67 LAPPVP 72

Search completed: April 6, 2006, 09:31:52
 Job time : 114.289 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 6, 2006, 09:35:13 ; Search time 12.7105 Seconds
(without alignments)
52.989 Million cell updates/sec

Title: US-10-632-388-296

Perfect score: 37

Sequence: 1 ILAPPVP 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Database :

PIR 80:*

1: PIR1:*

2: PIR2:*

3: PIR3:*

4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	37	100.0	1615	2 B49502	protein-tyrosine-p
2	37	100.0	1767	2 A49502	protein-tyrosine-p
3	35	94.6	266	2 F70595	hypothetical prote
4	35	94.6	441	2 T24591	hypothetical prote
5	34	91.9	189	2 D71658	3-octaprenyl-4-hyd
6	34	91.9	924	2 S75284	chemotaxis protein
7	33	89.2	113	2 AE1448	hypothetical prote
8	33	89.2	113	2 AE1729	hypothetical prote
9	33	89.2	113	2 AG1752	hypothetical prote
10	33	89.2	205	2 S41002	hypothetical prote
11	33	89.2	465	2 T35712	integral membrane
12	33	89.2	538	2 S76481	hypothetical prote
13	33	89.2	543	2 AF1927	hypothetical prote
14	33	89.2	551	2 T13994	envelope protein -
15	33	89.2	687	2 T08877	Modin - Podospora
16	33	89.2	819	2 F87708	cell division prot
17	33	89.2	981	2 A41401	mineralocorticoid
18	33	89.2	984	2 A29513	mineralocorticoid
19	33	89.2	1367	2 T21913	hypothetical prote
20	33	89.2	7962	2 I38346	elastic titin - hu
21	32	86.5	191	2 A97800	hypothetical prote
22	32	86.5	210	2 H72585	hypothetical prote
23	32	86.5	228	2 T10399	hypothetical prote
24	32	86.5	228	2 D30857	hypothetical prote
25	32	86.5	481	2 E86433	protein T17H7.3 [i
26	32	86.5	581	2 T24393	protein T17H7.3 [i
27	32	86.5	642	2 C86152	T7123-2 protein -
28	32	86.5	720	2 A96807	hypothetical prote
29	32	86.5	753	2 T01619	hypothetical prote

30	86.5	1069	2	T00377	KIAA0642 protein -
31	86.5	1610	2	D89451	protein T04G9.1 [i
32	86.5	1626	2	A75613	hypothetical prote
33	86.5	1630	2	T00390	KIAA0614 protein -
34	86.5	1791	2	T24089	hypothetical prote
35	86.5	3530	2	A59266	unconventional myo
36	86.5	3530	2	H87656	3-octaprenyl-4-hyd
37	86.5	162	2	H89164	protein T06E4.8 [i
38	86.5	198	2	H89164	hypothetical prote
39	86.5	199	2	T24595	hypothetical prote
40	86.5	205	2	T51884	hypothetical prote
41	86.5	210	2	G95858	probable 3-octapre
42	86.5	212	2	T24594	hypothetical prote
43	86.5	213	2	G99164	protein T06E4.9 [i
44	86.5	237	2	G95271	hypothetical prote
45	86.5	239	2	T42539	hypothetical prote
46	86.5	243	1	B42768	homeotic protein g
47	86.5	243	1	I51424	homeotic protein g
48	86.5	326	2	AF2138	hypothetical prote
49	86.5	332	2	T51269	hypothetical prote
50	86.5	347	2	A96835	gibberellin 3 beta
51	86.5	347	2	T51691	gibberellin 3 beta
52	86.5	384	2	H88924	protein F33E11.2 [
53	86.5	402	2	S77057	hypothetical prote
54	86.5	410	1	IXBE10	alpha trans-induci
55	86.5	474	2	I50830	Ig mu chain - Lepi
56	86.5	486	2	A41537	DNA-binding protei
57	86.5	565	2	AC3637	3-hydroxybutyryl-C
58	86.5	619	2	T28998	hypothetical prote
59	86.5	674	2	A97490	hypothetical prote
60	86.5	877	1	A25962	mitosis inhibitor
61	86.5	958	2	T13593	hypothetical prote
62	86.5	1027	2	T46481	hypothetical prote
63	86.5	1230	2	T18256	probable serine/th
64	86.5	1230	2	T18259	serine/threonine p
65	86.5	1306	2	T13592	hypothetical prote
66	86.5	1530	1	S01393	DNA-directed RNA p
67	86.5	1547	2	J00096	hypothetical 176K
68	86.5	1611	2	T38236	hypothetical prote
69	86.5	2606	2	T03159	large tegument pro
70	86.5	111	2	T49579	hypothetical prote
71	86.5	115	2	T50390	hypothetical prote
72	86.5	136	2	C72524	hypothetical prote
73	86.5	156	2	T43957	hypothetical prote
74	86.5	163	2	A70847	hypothetical prote
75	86.5	177	2	JC4839	delta Spi-B transc
76	86.5	179	2	A69104	hypothetical prote
77	86.5	194	2	E87504	hypothetical prote
78	86.5	197	2	E95411	probable decarboxy
79	86.5	203	2	AG3405	hypothetical exonu
80	86.5	212	2	H87472	hypothetical prote
81	86.5	231	2	T27803	hypothetical prote
82	86.5	250	2	D75317	hypothetical prote
83	86.5	262	2	S25655	Spi-B protein - hu
84	86.5	271	2	G83554	hypothetical prote
85	86.5	275	2	T24608	hypothetical prote
86	86.5	279	2	G75498	aminoglycoside N3-
87	86.5	281	2	T24658	hypothetical prote
88	86.5	296	2	G86446	unknown protein [i
89	86.5	304	2	T04053	nodulin-26 homolog
90	86.5	363	2	A41880	hypothetical prote
91	86.5	365	2	T04718	hypothetical prote
92	86.5	367	2	I39172	cyclin-dependent k
93	86.5	370	2	AB2490	hypothetical prote
94	86.5	375	2	T52606	squamosa promoter
95	86.5	395	2	A82349	hypothetical prote
96	86.5	402	2	J80282	cell division prot
97	86.5	431	2	AF2635	exopolysaccharide
98	86.5	435	2	E97417	exoQ-like protein
99	86.5	458	2	F86433	protein T17H7.5 [i
100	86.5	469	2	C87629	major facilitator
101	86.5	476	2	JC7304	fatty acid hydrolase
102	86.5	477	2	A47236	zinc-finger protei
103	86.5	480	2	S74228	fatty acid hydrolase

103	30	81.1	494	2	A42170	zinc finger protei	176	29	78.4	197	1	JH0309	tumor necrosis fac
104	30	81.1	497	2	J55076	myc-associated zin	177	29	78.4	197	2	S22010	pyrimidine synth
105	30	81.1	500	2	T14826	transcription init	178	29	78.4	197	2	T19892	hypothetical prote
106	30	81.1	509	2	T02864	probable Zn finger	179	29	78.4	204	2	T28010	hypothetical prote
107	30	81.1	510	2	A45338	connexin-56 - chic	180	29	78.4	215	2	T35493	hypothetical prote
108	30	81.1	541	1	S45428	PGT12 protein - y	181	29	78.4	216	2	T02658	probable germin pr
109	30	81.1	542	2	T32869	hypothetical prote	182	29	78.4	218	2	C86317	hypothetical prote
110	30	81.1	567	2	A62622	hypothetical prote	183	29	78.4	218	2	T26826	hypothetical prote
111	30	81.1	614	2	S58306	WD-40 repeat regul	184	29	78.4	224	2	C48652	transfer protein s
112	30	81.1	624	1	RDYCS7	sulfite reductase	185	29	78.4	242	2	S62017	phenylacrylic acid
113	30	81.1	634	2	T00359	hypothetical prote	186	29	78.4	243	2	D84792	probable protein t
114	30	81.1	641	2	T30076	hypothetical prote	187	29	78.4	245	2	T20831	hypothetical prote
115	30	81.1	643	2	E70682	probable gamma-glu	188	29	78.4	247	2	T26493	hypothetical prote
116	30	81.1	650	2	A11974	ferredoxin-sulfite	189	29	78.4	248	2	AH2339	hypothetical prote
117	30	81.1	708	2	T00064	hypothetical prote	190	29	78.4	250	2	A54308	PG-1 protein - num
118	30	81.1	722	2	G75563	phosphate acetyltr	191	29	78.4	255	2	T41451	very hypothetical
119	30	81.1	726	2	T35865	probable cyclase -	192	29	78.4	259	1	T22596	hypothetical prote
120	30	81.1	743	2	T13673	hypothetical prote	193	29	78.4	263	1	S23009	insulin-like growt
121	30	81.1	758	2	T15577	hypothetical prote	194	29	78.4	272	1	A36082	insulin-like growt
122	30	81.1	763	2	A10707	phosphatidylglycer	195	29	78.4	272	1	T48600	insulin-like growt
123	30	81.1	801	2	T52605	squamosa promoter	196	29	78.4	274	2	T04619	hypothetical prote
124	30	81.1	814	2	G02878	cadherin-15 precu	197	29	78.4	274	2	T47730	hypothetical prote
125	30	81.1	817	2	S51342	verprolin - yeast	198	29	78.4	278	2	A97471	hypothetical prote
126	30	81.1	898	2	H87481	ribonuclease, Rne/	199	29	78.4	278	2	AF2689	conserved hypothet
127	30	81.1	1008	2	T04462	hypothetical prote	200	29	78.4	283	2	D83948	hypothetical prote
128	30	81.1	1023	2	J40113	major acidic nucle	201	29	78.4	297	2	T14281	P40-like ribosomal
129	30	81.1	1064	2	T13963	formin related pro	202	29	78.4	304	1	XMECHD	acetyl-CoA carboxy
130	30	81.1	1069	2	D85383	hypothetical prote	203	29	78.4	304	2	A85873	acetyl-CoA carboxy
131	30	81.1	1072	2	T18802	hypothetical prote	204	29	78.4	304	2	H91028	acetyl-CoA carboxy
132	30	81.1	1085	2	T03531	cobN protein homol	205	29	78.4	309	2	E87559	peptidase M20/M25/
133	30	81.1	1237	2	A34598	ecdysone-induced p	206	29	78.4	317	2	S35282	eaA protein - phag
134	30	81.1	1245	2	H83574	conserved hypothet	207	29	78.4	323	2	S62085	hrpW protein - Pse
135	30	81.1	1401	2	T02255	probable ubiquitou	208	29	78.4	325	2	D70666	probable modD prot
136	30	81.1	1480	2	S48440	poly(A)-specific r	209	29	78.4	335	2	G64817	probable membrane
137	30	81.1	1538	2	E70874	probable ppsB prot	210	29	78.4	335	2	G85589	hypothetical prote
138	30	81.1	1557	2	D41214	protein-tyrosine-p	211	29	78.4	335	2	E90739	hypothetical prote
139	30	81.1	1630	2	C41214	protein-tyrosine-p	212	29	78.4	336	2	E47301	Vir10 homolog - Bo
140	30	81.1	1781	1	A34374	DNA-directed RNA p	213	29	78.4	336	2	T30459	hypothetical prote
141	30	81.1	3010	1	GNWVTC	genome polyprotein	214	29	78.4	345	2	AD3537	oligopeptide trans
142	30	81.1	3010	1	S18030	genome polyprotein	215	29	78.4	347	2	T13891	NADH2 dehydrogenas
143	29	78.4	17	2	A49237	45/47K antigen - M	216	29	78.4	348	2	C69399	N5,N10-methylene
144	29	78.4	29	2	C61233	conceptus protein	217	29	78.4	350	1	A45663	thymidine kinase (
145	29	78.4	61	2	A03152	hypothetical prote	218	29	78.4	353	1	C54258	transcription fact
146	29	78.4	67	2	G98135	hypothetical prote	219	29	78.4	354	1	S35090	transcription fact
147	29	78.4	92	2	S52277	hypothetical prote	220	29	78.4	358	2	T13017	hypothetical prote
148	29	78.4	108	2	G72474	hypothetical prote	221	29	78.4	363	2	T16755	hypothetical prote
149	29	78.4	115	2	T30042	hypothetical prote	222	29	78.4	365	2	S56792	hypothetical prote
150	29	78.4	122	2	T32932	hypothetical prote	223	29	78.4	370	2	C83120	probable RND efflu
151	29	78.4	124	2	S25113	insulin-like growt	224	29	78.4	371	2	T49696	related to peroxia
152	29	78.4	127	2	PC2265	cytochrome P450 pr	225	29	78.4	371	2	B39625	T-cell receptor al
153	29	78.4	127	2	B72732	hypothetical prote	226	29	78.4	372	2	S69574	hypothetical prote
154	29	78.4	128	2	T32947	hypothetical prote	227	29	78.4	373	1	VBETC	glycoprotein A - t
155	29	78.4	131	2	A84613	hypothetical prote	228	29	78.4	373	2	T16529	hypothetical prote
156	29	78.4	133	2	E70795	hypothetical prote	229	29	78.4	374	2	S48639	fructose-bisphosph
157	29	78.4	136	2	G75482	hypothetical prote	230	29	78.4	378	2	T12621	Dec3 promoter-bindi
158	29	78.4	138	2	E95917	hypothetical proli	231	29	78.4	386	2	S40256	glycine amidinotra
159	29	78.4	152	2	AH3007	conserved hypothet	232	29	78.4	386	2	F72773	probable molybdopt
160	29	78.4	162	2	E98276	hypothetical prote	233	29	78.4	397	2	A39565	lymphoid enhancer-
161	29	78.4	162	2	E87094	probable phosphope	234	29	78.4	398	2	A40966	porphobilinogen sy
162	29	78.4	162	2	H70865	probable isomerase	235	29	78.4	399	2	A39625	T-cell receptor al
163	29	78.4	163	2	F86438	protein t19E23.3 (236	29	78.4	402	2	D70908	probable lprL prot
164	29	78.4	164	2	E75293	hypothetical prote	237	29	78.4	407	2	C70816	hypothetical prote
165	29	78.4	172	2	A70864	hypothetical prote	238	29	78.4	409	2	T33019	hypothetical prote
166	29	78.4	185	2	D75602	hypothetical prote	239	29	78.4	411	2	A70509	hypothetical prote
167	29	78.4	189	1	XMECPD	3-octaprenyl-4-hyd	240	29	78.4	412	2	T26056	hypothetical prote
168	29	78.4	189	2	AB0801	probable decarboxy	241	29	78.4	413	2	AH2743	conserved hypothet
169	29	78.4	189	2	C91028	3-octaprenyl-4-hyd	242	29	78.4	414	2	I48975	apolipoprotein A-I
170	29	78.4	189	2	D85872	3-octaprenyl-4-hyd	243	29	78.4	414	2	A37113	apolipoprotein A-I
171	29	78.4	189	2	T05383	hypothetical prote	244	29	78.4	415	2	T34156	hypothetical prote
172	29	78.4	190	2	T35570	hypothetical prote	245	29	78.4	415	2	T26057	hypothetical prote
173	29	78.4	196	2	D98190	decarboxylase [imp	246	29	78.4	416	2	G97524	hypothetical prote
174	29	78.4	196	2	AE3095	phenylacrylic acid	247	29	78.4	420	2	B75333	twitching mobility
175	29	78.4	196	2	AC3629	3-octaprenyl-4-hyd	248	29	78.4	422	2	G70984	probable papA5 pro

249	29	78.4	423	2	S73020	hypothetical prote	322	29	78.4	762	2	T14815	hypothetical prote
250	29	78.4	423	2	T18784	hypothetical prote	323	29	78.4	766	2	S59657	hypothetical prote
251	29	78.4	424	2	A50000	porphobilinogen sy	324	29	78.4	771	2	S56791	hypothetical prote
252	29	78.4	430	2	S58169	porphobilinogen sy	325	29	78.4	798	2	F96714	probable protease
253	29	78.4	430	2	D96719	hypothetical prote	326	29	78.4	802	2	H59434	oligophrenin 1, Rn
254	29	78.4	447	1	A70773	diaminopimelate de	327	29	78.4	810	2	C70791	probable ponA' pro
255	29	78.4	458	2	T06313	hypothetical prote	328	29	78.4	832	2	F98128	insertion element
256	29	78.4	464	2	T36256	probable membrane	329	29	78.4	844	2	S61104	BR01 protein - yea
257	29	78.4	465	2	S68987	transcription acti	330	29	78.4	848	2	I55498	testicular dynamin
258	29	78.4	467	2	A59268	modulation competi	331	29	78.4	851	2	B40671	dynamin, internal
259	29	78.4	468	2	S26741	T-cell glycoprotei	332	29	78.4	851	2	S11508	D100 protein - rat
260	29	78.4	472	2	S39507	glucuronosyl trans	333	29	78.4	864	2	A40671	dynamin, internal
261	29	78.4	478	1	S03826	UMP synthase - eli	334	29	78.4	868	2	A30817	serendipity (sry h
262	29	78.4	482	2	T36045	hypothetical prote	335	29	78.4	868	2	AT3204	autotransporter pr
263	29	78.4	486	2	AG2199	hypothetical prote	336	29	78.4	877	2	S49197	envelope protein p
264	29	78.4	488	2	F96724	hypothetical prote	337	29	78.4	896	2	S36326	clathrin assembly
265	29	78.4	489	1	VGBETA	glycoprotein A - t	338	29	78.4	901	2	A44825	phosphoprotein, sy
266	29	78.4	494	2	T39763	mitochondrial proc	339	29	78.4	903	2	T00074	hypothetical prote
267	29	78.4	495	2	A40465	alanine transamina	340	29	78.4	915	2	S36327	clathrin assembly
268	29	78.4	495	2	A39900	alanine transamina	341	29	78.4	939	2	C70876	hypothetical prote
269	29	78.4	501	1	A60005	glycoprotein A pre	342	29	78.4	955	2	F84914	hypothetical prote
270	29	78.4	501	1	VGBEMA	glycoprotein A pre	343	29	78.4	963	2	T19140	hypothetical prote
271	29	78.4	501	1	VGBEMB	glycoprotein A pre	344	29	78.4	992	2	G70950	hypothetical prote
272	29	78.4	502	2	T00483	hypothetical prote	345	29	78.4	1011	2	T07712	probable ABC-type
273	29	78.4	503	2	S36549	L1 protein - human	346	29	78.4	1013	2	T33470	hypothetical prote
274	29	78.4	505	1	VGBEMH	glycoprotein gp57-	347	29	78.4	1018	2	T22318	hypothetical prote
275	29	78.4	505	2	D69382	conserved hypothet	348	29	78.4	1023	2	B59430	KIAA0189 protein [
276	29	78.4	509	2	T29291	hypothetical prote	349	29	78.4	1048	2	T31425	C-terminal domain-
277	29	78.4	514	2	C49507	potassium channel	350	29	78.4	1087	2	T49496	hypothetical prote
278	29	78.4	515	2	B83477	hypothetical prote	351	29	78.4	1114	2	T49517	p63 related protei
279	29	78.4	517	2	A48250	u2AF-homologous pr	352	29	78.4	1140	2	D88690	protein F41H10.3 [
280	29	78.4	520	2	T26201	hypothetical prote	353	29	78.4	1175	2	T20346	pyruvate carboxyla
281	29	78.4	523	1	A60408	glycoprotein gp57-	354	29	78.4	1186	2	C64588	cag pathogenicity
282	29	78.4	525	2	T20662	hypothetical prote	355	29	78.4	1193	2	JC2489	peptidyl-di-peptida
283	29	78.4	541	2	E87414	transglycosylase,	356	29	78.4	1213	2	A41724	limb deformity (ld
284	29	78.4	548	2	T23270	acetyl choline rec	357	29	78.4	1218	2	S71376	glutamate receptor
285	29	78.4	548	2	S59133	ETS2 repressor fac	358	29	78.4	1233	2	T10365	helicase - Orgyia
286	29	78.4	553	1	GLBP5V	gene 1 protein - s	359	29	78.4	1232	2	S40766	hypothetical prote
287	29	78.4	562	2	AD0304	probable hemolysin	360	29	78.4	1236	2	E70977	hypothetical prote
288	29	78.4	562	2	AD0074	probable hemolysin	361	29	78.4	1297	2	S25714	son-of-sevenless-2
289	29	78.4	574	2	T43556	Wiskott-Aldrich sy	362	29	78.4	1306	1	A31759	peptidyl-di-peptida
290	29	78.4	574	2	T39819	wiskott-aldrich sy	363	29	78.4	1307	2	G96711	unknown protein, 9
291	29	78.4	575	2	S59740	NRD1 protein - yea	364	29	78.4	1307	2	T17453	ERG-associated pro
292	29	78.4	579	2	E75275	hypothetical prote	365	29	78.4	1312	1	A34171	peptidyl-di-peptida
293	29	78.4	590	2	S16411	terminase Arpase c	366	29	78.4	1313	1	JC2038	peptidyl-di-peptida
294	29	78.4	602	2	A49507	potassium channel	367	29	78.4	1360	2	T34302	cell polarity prot
295	29	78.4	605	2	T27397	hypothetical prote	368	29	78.4	1367	2	T33819	hypothetical prote
296	29	78.4	607	2	T34040	hypothetical prote	369	29	78.4	1413	2	T36467	hypothetical prote
297	29	78.4	609	2	T01892	hypothetical prote	370	29	78.4	1417	2	A57570	Bloom's syndrome r
298	29	78.4	611	2	T21747	hypothetical prote	371	29	78.4	1457	2	T14577	protein kinase Yak
299	29	78.4	618	2	B87564	conserved hypothet	372	29	78.4	1469	2	T09219	basal transcriptio
300	29	78.4	621	2	G83474	hypothetical prote	373	29	78.4	1490	2	S72351	nonstructural poly
301	29	78.4	623	2	G70644	probable spA prot	374	29	78.4	1533	2	A46221	abdominal segment
302	29	78.4	625	2	S69707	hypothetical prote	375	29	78.4	1634	2	T26517	hypothetical prote
303	29	78.4	628	2	A84087	hypothetical prote	376	29	78.4	1634	2	T36467	hypothetical prote
304	29	78.4	636	2	T38010	hypothetical prote	377	29	78.4	1872	2	T48818	glucan 1,4-alpha-9
305	29	78.4	644	2	A53184	myc far upstream e	378	29	78.4	2150	2	T08165	RNAi polyprotein -
306	29	78.4	652	2	T04856	hypothetical prote	379	29	78.4	2225	1	A23443	pyrimidine synthe
307	29	78.4	678	2	A75580	hypothetical prote	380	29	78.4	2240	2	T37057	probable multi-dom
308	29	78.4	684	2	T25603	hypothetical prote	381	29	78.4	2297	2	T34918	polyketide syntha
309	29	78.4	685	2	S64158	hypothetical prote	382	29	78.4	2344	2	T41590	probable sensor-li
310	29	78.4	693	2	A41617	dopamine transport	383	29	78.4	2493	2	S72349	nonstructural poly
311	29	78.4	717	1	S29923	transforming prote	384	29	78.4	2493	2	S26372	polyprotein - marm
312	29	78.4	723	2	B55926	DNA binding protei	385	29	78.4	2970	2	T08839	polyprotein - dour
313	29	78.4	731	2	T19721	hypothetical prote	386	29	78.4	3005	2	T08841	genome polyprotein
314	29	78.4	732	1	A35655	peptidyl-di-peptida	387	29	78.4	3010	1	GNWVTW	giant protein p619
315	29	78.4	732	1	S05238	peptidyl-di-peptida	388	29	78.4	4861	2	S71752	hypothetical prote
316	29	78.4	737	2	S28030	DNA-binding protei	389	29	78.4	13055	2	T16580	ribosomal protein
317	29	78.4	743	2	C56695	transducin-like en	390	28	75.7	12	2	PA0098	hypothetical prote
318	29	78.4	745	2	S13586	triacylglycerol li	391	28	75.7	62	2	G81143	hypothetical prote
319	29	78.4	750	2	D86245	hypothetical prote	392	28	75.7	62	2	T36447	hypothetical prote
320	29	78.4	750	2	AH3158	hypothetical prote	393	28	75.7	63	2	F83293	hypothetical prote
321	29	78.4	751	2	T34490	hypothetical prote	394	28	75.7	90	2	D24356	hypothetical prote

395	28	75.7	92	2	S49385	movement protein V	468	28	75.7	336	2	AG0307	conserved hypothet
396	28	75.7	101	2	H75370	stage V sporulatio	469	28	75.7	339	2	E83419	probable permease
397	28	75.7	105	2	H72736	hypothetical prote	470	28	75.7	340	1	QBES6	UL20 protein precu
398	28	75.7	107	1	WMBEL2	latency-related pr	471	28	75.7	343	2	S09272	ig alpha chain C r
399	28	75.7	117	2	T26804	hypothetical prote	472	28	75.7	345	2	B75274	conserved hypothet
400	28	75.7	118	2	T26804	hypothetical prote	473	28	75.7	348	2	D88088	protein B0454.1 fi
401	28	75.7	121	2	S65808	transcription acti	474	28	75.7	350	2	B82713	lipase modulator X
402	28	75.7	124	2	T25799	hypothetical prote	475	28	75.7	356	2	T51944	pathogenicity MAP
403	28	75.7	127	2	FC2266	cytochrome P450 pr	476	28	75.7	357	2	S09265	ig alpha chain C r
404	28	75.7	130	2	B72469	hypothetical prote	477	28	75.7	359	1	K1BBBT	thymidine kinase (
405	28	75.7	138	2	T35301	hypothetical prote	478	28	75.7	361	2	T51943	mitogen-activated
406	28	75.7	141	2	T19578	hypothetical prote	479	28	75.7	361	2	J02379	homeobox 1 protein
407	28	75.7	151	2	D87663	hypothetical prote	480	28	75.7	362	2	H82178	conserved hypothet
408	28	75.7	153	2	T52498	hypothetical prote	481	28	75.7	369	2	E81722	hypothetical prote
409	28	75.7	158	2	S46257	mitochondrial inne	482	28	75.7	374	2	G84596	hypothetical prote
410	28	75.7	160	2	S09823	hypothetical prote	483	28	75.7	383	2	H75263	probable zinc prot
411	28	75.7	164	2	T26608	hypothetical prote	484	28	75.7	396	2	S22909	interleukin-3 rece
412	28	75.7	174	2	B72239	conserved hypothet	485	28	75.7	403	2	S52796	prpL2 protein - hu
413	28	75.7	184	2	F64312	phenylacrylic acid	486	28	75.7	407	2	T00693	hypothetical prote
414	28	75.7	185	2	T00519	proline-rich prote	487	28	75.7	418	2	A37344	hypothetical prote
415	28	75.7	186	2	T45838	hypothetical prote	488	28	75.7	420	2	A55283	acrosin (EC 3.4.21
416	28	75.7	190	2	T26609	hypothetical prote	489	28	75.7	426	2	JC7773	IL-13ralpha 1 prot
417	28	75.7	200	2	B95349	Probable decarboxy	490	28	75.7	428	2	AH3536	dihydroliipoamide S
418	28	75.7	202	2	JN0579	DNA-binding protei	491	28	75.7	433	2	T36122	probable integral
419	28	75.7	202	2	T01376	hypothetical prote	492	28	75.7	433	2	A10536	probable membrane
420	28	75.7	204	2	T31290	3-octaprenyl-4-hyd	493	28	75.7	434	2	S62168	hypothetical prote
421	28	75.7	207	2	C86147	nitrate/nitrite re	494	28	75.7	436	2	JX0172	acrosin (EC 3.4.21
422	28	75.7	226	2	S77174	probable hypoesen	495	28	75.7	436	2	JQ1459	Bt1 protein precu
423	28	75.7	227	2	D85018	thrombospondin 1 -	496	28	75.7	439	2	S58327	cobalt accumulatio
424	28	75.7	229	2	S57957	hypothetical prote	497	28	75.7	448	2	JN0118	glucan 1,3-beta-gl
425	28	75.7	244	2	T37686	hypothetical prote	498	28	75.7	449	2	T02625	hypothetical prote
426	28	75.7	245	2	S63455	hypothetical prote	499	28	75.7	452	2	JC4100	hydroxyindole O-me
427	28	75.7	252	2	I46859	MHC RLA - rabbit (500	28	75.7	452	2	S46798	hypothetical prote
428	28	75.7	256	2	H96524	protein TiN15.11 (501	28	75.7	469	2	AE2188	hypothetical prote
429	28	75.7	257	2	B84860	hypothetical prote	502	28	75.7	472	2	T03169	probable glycoprot
430	28	75.7	259	2	T36003	hypothetical prote	503	28	75.7	473	1	RGBYM3	regulatory protein
431	28	75.7	260	2	G87349	conserved hypothet	504	28	75.7	478	2	I53960	PR2 alpha - human
432	28	75.7	262	2	A38114	resolvase rsd - Sa	505	28	75.7	489	2	B33416	nuclear factor I -
433	28	75.7	262	2	T01562	hypothetical prote	506	28	75.7	491	2	T07598	proline-rich prote
434	28	75.7	262	2	T48954	hypothetical prote	507	28	75.7	494	2	AC1742	protein involved i
435	28	75.7	263	2	D71342	hypothetical prote	508	28	75.7	494	2	AB1372	proteins involved
436	28	75.7	268	2	A24716	D protein - Escher	509	28	75.7	497	2	T27012	hypothetical prote
437	28	75.7	268	2	T00239	resolvase - Escher	510	28	75.7	498	2	F89861	Na+/H+-antiporter
438	28	75.7	269	2	T28957	hypothetical prote	511	28	75.7	499	2	S01038	transcription fact
439	28	75.7	271	2	T40982	hypothetical prote	512	28	75.7	504	2	G02474	interferon regulat
440	28	75.7	275	2	G75606	probable glycohydr	513	28	75.7	506	2	S37156	NADPH-ferrihemop
441	28	75.7	281	2	D70845	hypothetical prote	514	28	75.7	506	2	A33416	nuclear factor I -
442	28	75.7	283	2	E88597	protein Y47D3B.6 (515	28	75.7	510	2	S45571	nuclear factor I-C
443	28	75.7	297	2	S21737	proline-rich prote	516	28	75.7	533	1	TVFVFP	protein-tyrosine k
444	28	75.7	299	2	S60521	envelope polyprote	517	28	75.7	538	2	I88093	PRR2 delta - human
445	28	75.7	299	2	S60552	envelope polyprote	518	28	75.7	538	2	T22654	hypothetical prote
446	28	75.7	299	2	S60551	envelope polyprote	519	28	75.7	547	2	C96828	unknown protein F1
447	28	75.7	299	2	S60523	envelope polyprote	520	28	75.7	547	2	S77410	translation releas
448	28	75.7	299	2	S60554	envelope polyprote	521	28	75.7	554	2	S75016	phosphoglycerate d
449	28	75.7	299	2	S60553	envelope polyprote	522	28	75.7	554	2	F86244	hypothetical prote
450	28	75.7	300	1	S57749	SURF1 protein - hu	523	28	75.7	557	2	T41495	hypothetical prote
451	28	75.7	300	2	AG0983	dipeptide transpor	524	28	75.7	560	2	F84341	hypothetical prote
452	28	75.7	300	2	S60522	envelope polyprote	525	28	75.7	572	2	H96685	probable AMP-bindi
453	28	75.7	302	2	T45937	hypothetical prote	526	28	75.7	574	2	T05964	probable low-affin
454	28	75.7	303	2	S60549	envelope polyprote	527	28	75.7	575	2	B71525	hypothetical prote
455	28	75.7	303	2	S60550	envelope polyprote	528	28	75.7	576	2	T18067	hypothetical prote
456	28	75.7	304	2	F75278	conserved hypothet	529	28	75.7	582	2	S57722	sl1-1 protein - Ca
457	28	75.7	305	2	T19710	hypothetical prote	530	28	75.7	585	2	T19814	hypothetical prote
458	28	75.7	310	2	T33611	hypothetical prote	531	28	75.7	588	2	AE1926	two-component resp
459	28	75.7	312	2	A56911	TRADD protein - hu	532	28	75.7	589	2	T42244	probable membrane-
460	28	75.7	315	2	A83496	conserved hypothet	533	28	75.7	598	2	E72733	hypothetical prote
461	28	75.7	316	2	D75588	transcription regu	534	28	75.7	600	2	T24447	hypothetical prote
462	28	75.7	316	2	T34023	hypothetical prote	535	28	75.7	621	2	T24090	hypothetical prote
463	28	75.7	321	2	T03431	probable transport	536	28	75.7	631	2	T05016	hypothetical prote
464	28	75.7	321	2	AB3245	hypothetical prote	537	28	75.7	634	2	S33339	sonatotropin recep
465	28	75.7	324	2	E83122	probable iron/asco	538	28	75.7	635	2	F75477	hypothetical prote
466	28	75.7	335	2	S09275	ig alpha chain C r	539	28	75.7	636	2	A84903	hypothetical prote
467	28	75.7	336	2	A83166	conserved hypothet	540	28	75.7	638	2	A33991	sonatotropin recep

541	28	75.7	638	2	S12136	somatotropin recep	614	28	75.7	1529	2	T16779	hypothetical prote
542	28	75.7	638	2	B28176	somatotropin recep	615	28	75.7	1599	2	T15854	hypothetical prote
543	28	75.7	638	2	A33505	somatotropin recep	616	28	75.7	1733	2	S27939	tensin - chicken
544	28	75.7	643	2	S60710	alpha 1,2 mannosid	617	28	75.7	1744	2	A54970	tensin, cardiac mu
545	28	75.7	643	2	T23453	hypothetical prote	618	28	75.7	1792	2	A57075	tensin - chicken (
546	28	75.7	646	1	W2BEC8	68.6K capsid prote	619	28	75.7	1824	2	T07589	disease resistance
547	28	75.7	677	2	S33608	somatotropin-bindi	620	28	75.7	1976	2	I56555	sodium channel pro
548	28	75.7	699	2	C43674	US4 protein - huma	621	28	75.7	2133	2	T42763	coagulation factor
549	28	75.7	704	2	T24517	hypothetical prote	622	28	75.7	2403	2	A59386	sanko - human
550	28	75.7	705	2	T10723	NADPH-ferrihemopro	623	28	75.7	2411	2	A46299	cytosine kinase su
551	28	75.7	712	2	T31523	hypothetical prote	624	28	75.7	2531	2	T31070	notch homolog - se
552	28	75.7	717	2	T27066	hypothetical prote	625	28	75.7	3010	1	GNWVCJ	genome polyprotein
553	28	75.7	738	2	S14270	alcohol dehydrogen	626	28	75.7	3157	2	B70969	probable PPE prote
554	28	75.7	770	2	S76883	nitric-oxide reduc	627	28	75.7	3624	2	AD0835	large repetitive p
555	28	75.7	780	2	T27941	hypothetical prote	628	28	75.7	3869	2	A48205	All-1 protein +GPE
556	28	75.7	786	2	E87565	conserved hypothet	629	28	75.7	4936	2	AH2515	hypothetical prote
557	28	75.7	791	2	A53691	diacylglycerol kin	630	28	75.7	4957	2	T03455	ALR protein - huma
558	28	75.7	793	2	S59067	penton long fiber	631	28	75.7	5262	2	T03454	ALR protein - huma
559	28	75.7	794	2	B87605	sensory box histid	632	27	73.0	57	2	S10782	salivary protein P
560	28	75.7	820	2	T00645	hypothetical prote	633	27	73.0	64	2	B41286	granulocyte-macrop
561	28	75.7	824	2	I50618	c-fps proto oncoge	634	27	73.0	68	2	E85918	hypothetical prote
562	28	75.7	824	2	T23923	hypothetical prote	635	27	73.0	79	1	PUHUSB	proline-rich pepti
563	28	75.7	834	2	JC8035	N-acetyltransferas	636	27	73.0	87	2	T00171	hypothetical prote
564	28	75.7	845	2	T00071	hypothetical prote	637	27	73.0	90	2	T18068	hypothetical prote
565	28	75.7	859	2	S51646	inscuteable protei	638	27	73.0	107	2	C72623	hypothetical prote
566	28	75.7	865	2	A47282	calcium-binding pr	639	27	73.0	109	2	S64309	hypothetical prote
567	28	75.7	866	1	JC4305	dynamain II - human	640	27	73.0	113	2	AE0176	conserved hypothet
568	28	75.7	868	2	A36878	dynamain 2 - rat	641	27	73.0	115	2	S75426	hypothetical prote
569	28	75.7	869	2	A25945	coagulation factor	642	27	73.0	116	2	A72685	hypothetical prote
570	28	75.7	870	2	B53165	dynamain II isoform	643	27	73.0	118	2	S28201	alpha-amylase inhi
571	28	75.7	870	2	A53165	dynamain II isoform	644	27	73.0	126	2	PC2273	cytochrome P450 pr
572	28	75.7	872	2	S33015	hypothetical prote	645	27	73.0	128	2	PC2260	cytochrome P450 pr
573	28	75.7	873	1	TVFVFS	protein-tyrosine k	646	27	73.0	130	2	S66474	cytochrome P450 C
574	28	75.7	873	1	TVFVFS	protein-tyrosine k	647	27	73.0	130	2	S66474	cytochrome P450 C
575	28	75.7	875	1	URXLA2	peptidylglycine mo	648	27	73.0	130	2	S66301	cytochrome P450 C
576	28	75.7	885	2	S67660	hypothetical prote	649	27	73.0	131	2	S66475	cytochrome P450 C
577	28	75.7	906	2	T00039	hypothetical prote	650	27	73.0	132	2	T11239	hypothetical prote
578	28	75.7	922	2	T37256	metalloproteinase	651	27	73.0	134	2	JC5572	proline-rich prote
579	28	75.7	927	2	A48085	transcription fact	652	27	73.0	137	2	S00696	vitellogenin precu
580	28	75.7	935	2	S17855	peptidylglycine mo	653	27	73.0	138	2	D96715	protein F4N2.10 [i
581	28	75.7	938	1	Q0B824	nuclear antigen BB	654	27	73.0	138	2	C84491	hypothetical prote
582	28	75.7	970	2	JQ0302	hypothetical 112K	655	27	73.0	141	2	S35611	cytochrome P450 4A
583	28	75.7	992	2	A31666	hypothetical prote	656	27	73.0	147	2	A36790	hypothetical prote
584	28	75.7	997	2	S44457	period clock prote	657	27	73.0	148	2	T02305	hypothetical prote
585	28	75.7	999	2	A34307	Ca2+-transporting	658	27	73.0	149	2	E95201	hypothetical prote
586	28	75.7	999	2	S72267	hypothetical prote	659	27	73.0	149	2	E98068	hypothetical prote
587	28	75.7	1013	2	T46422	hypothetical prote	660	27	73.0	152	2	JC4894	ubiquitin-protein
588	28	75.7	1029	2	H86179	hypothetical prote	661	27	73.0	159	2	T19624	hypothetical prote
589	28	75.7	1040	2	T29092	TSC-22 protein hom	662	27	73.0	162	2	AB0459	probable regulator
590	28	75.7	1050	2	JW0092	serine-threonine k	663	27	73.0	163	2	A71469	hypothetical prote
591	28	75.7	1053	2	S58883	calcium-channel ho	664	27	73.0	166	2	D87664	hypothetical prote
592	28	75.7	1087	2	C84263	transmembrane olig	665	27	73.0	169	2	T49618	2',3'-cyclic-nucle
593	28	75.7	1104	2	I38869	transcription fact	666	27	73.0	170	2	H75279	phenylacetic acid
594	28	75.7	1106	2	T44598	hypothetical prote	667	27	73.0	174	2	T45335	hypothetical prote
595	28	75.7	1106	2	T13938	gene shuttle craft	668	27	73.0	178	2	I50697	non-collagenous al
596	28	75.7	1110	2	T49091	gamma response 1 p	669	27	73.0	181	2	S37543	H+-transporting tw
597	28	75.7	1114	2	T30299	dynamain heavy chain	670	27	73.0	181	2	T04643	hypothetical prote
598	28	75.7	1116	2	T38073	serine/threonine-p	671	27	73.0	181	2	T45581	hypothetical prote
599	28	75.7	1120	2	H88449	protein F54D8.1 [i	672	27	73.0	184	2	S48035	hypothetical prote
600	28	75.7	1166	2	T15628	hypothetical prote	673	27	73.0	187	2	H81347	probable decarboxy
601	28	75.7	1170	1	TSRUP1	thrombospondin 1 p	674	27	73.0	189	2	H81968	probable decarboxy
602	28	75.7	1206	2	S24407	formin isoform IV	675	27	73.0	190	2	B81023	3-octaprenyl-4-hyd
603	28	75.7	1240	2	T03097	CDO protein - huma	676	27	73.0	192	2	A55924	GTP-binding protei
604	28	75.7	1333	2	A37488	Ras guanine nucleo	677	27	73.0	195	2	T29395	hypothetical prote
605	28	75.7	1336	2	S25716	Ras guanine nucleo	678	27	73.0	197	2	F75436	conserved hypothet
606	28	75.7	1338	2	T30565	MAP kinase kinase	679	27	73.0	199	2	T16113	hypothetical prote
607	28	75.7	1409	2	T37188	presynaptic activi	680	27	73.0	203	2	G37380	phnH protein limpo
608	28	75.7	1464	2	T13716	bazooka gene prote	681	27	73.0	203	2	AF2598	conserved hypothet
609	28	75.7	1468	2	S11515	formin - mouse	682	27	73.0	204	2	G69761	phenylacrylic acid
610	28	75.7	1478	2	S20117	protein kinase BCK	683	27	73.0	205	2	S37804	hypothetical prote
611	28	75.7	1495	2	S60255	transcription co-r	684	27	73.0	216	2	S58652	hypothetical prote
612	28	75.7	1499	2	B59431	Rho GTPase activat	685	27	73.0	217	2	T24867	hypothetical prote
613	28	75.7	1503	2	T43166	alpha-2-macroglobu	686	27	73.0	218	2	E70944	probable lppI prot

687 27 73.0 221 2 D86600 SET domain protein
688 27 73.0 221 2 B81514 conserved hypothet
689 27 73.0 221 2 D74024 set domain protein
690 27 73.0 225 2 S55720 ribosomal protein
691 27 73.0 228 2 C70623 hypothetical prote
692 27 73.0 231 2 C70510 hypothetical prote
693 27 73.0 233 2 D86603 muramidase [import
694 27 73.0 233 2 A72022 muramidase (invasi
695 27 73.0 233 2 D89114 hypothetical prote
696 27 73.0 235 2 E84313 hypothetical prote
697 27 73.0 237 2 B70529 hypothetical prote
698 27 73.0 237 2 C72676 hypothetical prote
699 27 73.0 238 2 T11223 hypothetical prote
700 27 73.0 240 2 S53403 probable membrane
701 27 73.0 250 2 T36428 hypothetical prote
702 27 73.0 252 2 AD2558 hypothetical prote
703 27 73.0 252 2 T04739 hypothetical prote
704 27 73.0 254 2 AE2516 hypothetical prote
705 27 73.0 268 2 T15169 hypothetical prote
706 27 73.0 269 2 D88130 protein F10G7.6 [i
707 27 73.0 278 2 E82965 hypothetical prote
708 27 73.0 282 2 D95892 probable glycerol-
709 27 73.0 283 2 G69844 conserved hypothet
710 27 73.0 289 2 S74632 hypothetical prote
711 27 73.0 293 2 T45538 moaC protein [impo
712 27 73.0 294 1 S73832 MG237 homolog F10_
713 27 73.0 296 1 PDBYC REP2 protein - yea
714 27 73.0 296 2 T47598 hypothetical prote
715 27 73.0 299 2 QJ1930 V protein - rinder
716 27 73.0 300 2 AH2955 hypothetical prote
717 27 73.0 300 2 E98327 hypothetical prote
718 27 73.0 301 2 E72289 oligopeptide ABC t
719 27 73.0 302 2 H96792 unknown protein FI
720 27 73.0 303 2 T46416 hydrolase, probabl
721 27 73.0 307 2 B87659 coatomer complex e
722 27 73.0 308 2 I46019 coatomer complex e
723 27 73.0 308 2 T48080 hypothetical prote
724 27 73.0 311 2 F86341 homeotic protein -
725 27 73.0 316 2 S16681 probable prolina-r
726 27 73.0 321 2 F84597 lactose-binding le
727 27 73.0 324 2 A46631 conserved hypothet
728 27 73.0 325 2 D75555 lutropin-choriogon
729 27 73.0 329 2 B41344 lutropin-choriogon
730 27 73.0 329 2 D41344 zinc-finger protei
731 27 73.0 331 2 B47236 lutropin-choriogon
732 27 73.0 331 2 C41344 hypothetical prote
733 27 73.0 333 2 T23840 hypothetical prote
734 27 73.0 336 2 S29507 metavinuculin - pig
735 27 73.0 342 2 B84220 hypothetical prote
736 27 73.0 344 2 A86294 hypothetical prote
737 27 73.0 349 2 D88431 protein M88.4 [imp
738 27 73.0 350 2 E71248 hypothetical prote
739 27 73.0 352 2 T06482 probable cell wall
740 27 73.0 352 2 C96643 hypothetical prote
741 27 73.0 356 2 I40212 cytochrome P450 BJ
742 27 73.0 362 2 C90674 hypothetical prote
743 27 73.0 362 2 C64758 yaha protein - bac
744 27 73.0 363 2 B45900 complement C3d/Eps
745 27 73.0 365 2 H85524 hypothetical prote
746 27 73.0 366 2 C70907 hypothetical prote
747 27 73.0 367 2 H75440 alanine racemase -
748 27 73.0 368 2 T03694 cysteine proteinas
749 27 73.0 370 1 S55333 durrp kinase pim
750 27 73.0 371 1 WMBEY0 durrp diphosphatase
751 27 73.0 379 2 S31719 proline-rich prote
752 27 73.0 382 2 I39068 proline- arginine-
753 27 73.0 383 2 S76334 hypothetical prote
754 27 73.0 384 2 H70580 hypothetical prote
755 27 73.0 388 2 T27645 hypothetical prote
756 27 73.0 390 2 JE0212 hypothetical 44k p
757 27 73.0 390 2 JE0211 hypothetical 44k p
758 27 73.0 390 2 T49705 related to transcr
759 27 73.0 390 2 G01936 Abl binding protei

760 27 73.0 391 2 S54161
761 27 73.0 394 2 B75404
762 27 73.0 395 2 G87481
763 27 73.0 396 1 S33310
764 27 73.0 398 2 D83214
765 27 73.0 400 2 C70837 hypothetical prote
766 27 73.0 400 2 S06945 granulocyte-macrop
767 27 73.0 403 2 A53662 homeotic protein H
768 27 73.0 405 2 T10630 ethylene-regulated
769 27 73.0 409 2 S70704 carbon catabolite
770 27 73.0 410 2 B35670 protein-tyrosine k
771 27 73.0 412 2 T40899 probable proteinas
772 27 73.0 415 1 A34170 acrosin [EC 3.4.21
773 27 73.0 415 2 S12357 interleukin-5 rece
774 27 73.0 416 2 T28756 hypothetical prote
775 27 73.0 416 2 T38428 T-complex protein
776 27 73.0 416 2 A43561 homeotic protein m
777 27 73.0 417 2 T33376 hypothetical prote
778 27 73.0 419 2 T19871 hypothetical prote
779 27 73.0 420 2 H84182 hypothetical prote
780 27 73.0 422 2 S41734 glycine amidinotra
781 27 73.0 427 2 C83464 hypothetical prote
782 27 73.0 427 2 T23954 hypothetical prote
783 27 73.0 431 2 S47538 acrosin [EC 3.4.21
784 27 73.0 436 2 T22253 hypothetical prote
785 27 73.0 440 2 JC7807 Miskott-Aldrich sy
786 27 73.0 448 2 T06076 proline-rich prote
787 27 73.0 451 2 B70805 hypothetical prote
788 27 73.0 459 2 G86264 P319 hypothetical
789 27 73.0 461 2 H84644 hypothetical prote
790 27 73.0 463 2 T01872 hypothetical prote
791 27 73.0 465 2 G02738 FRAC-4 - human
792 27 73.0 466 1 P2WLJ1 L2 protein - human
793 27 73.0 469 2 A12833 UDP-N-acetylmuram
794 27 73.0 469 2 E97611 hypothetical prote
795 27 73.0 470 2 S28303 hypothetical prote
796 27 73.0 472 1 P2WL58 L2 protein - human
797 27 73.0 473 1 P2WLHS minor capaid prote
798 27 73.0 478 2 AG1051 probable transport
799 27 73.0 478 2 T03548 mannitol 2-dehydro
800 27 73.0 479 2 T09365 cytochrome P450 ho
801 27 73.0 480 2 T48311 cytochrome P450 52
802 27 73.0 480 2 S52306 zinc finger protei
803 27 73.0 480 2 C86357 hypothetical prote
804 27 73.0 481 2 T16484 hypothetical prote
805 27 73.0 483 2 T06712 probable cytochrom
806 27 73.0 484 2 D65230 hypothetical 52.9
807 27 73.0 484 2 A98275 hypothetical prote
808 27 73.0 484 2 A86116 hypothetical prote
809 27 73.0 487 2 B70720 hypothetical prote
810 27 73.0 488 2 T42038 catalase [EC 1.11.
811 27 73.0 490 2 T00404 probable cytochrom
812 27 73.0 490 2 AE3433 mitochondrial proc
813 27 73.0 491 2 D95962 probable transcrip
814 27 73.0 492 2 T43064 CEP-6 protein - Ca
815 27 73.0 496 1 S41192 cytochrome P450 4D
816 27 73.0 498 2 P84667 probable cytochrom
817 27 73.0 499 1 A40092 protein-tyrosine k
818 27 73.0 499 2 T49064 cytochrome P450-1i
819 27 73.0 501 2 S76563 hypothetical prote
820 27 73.0 501 2 T28970 hypothetical prote
821 27 73.0 503 1 S66374 cytochrome P450 4M
822 27 73.0 503 2 T02357 cytochrome P450 ho
823 27 73.0 503 2 T48825 hypothetical prote
824 27 73.0 504 1 S23558 ubiquinol-cytochro
825 27 73.0 506 1 A40164 cytochrome P450 4B
826 27 73.0 507 1 A32966 cytochrome P450 4A
827 27 73.0 507 1 A43387 polymerase-aseocia
828 27 73.0 507 2 JC7883 cytochrome P450 4X
829 27 73.0 507 2 JQ1929 phosphoprotein - r
830 27 73.0 507 2 C71412 probable hydroxypr
831 27 73.0 509 1 O4RTLQ laurate omega-hydr
832 27 73.0 509 2 S47553 cytochrome P450 Cy

833	27	73.0	510	1	A29368	prostaglandin omeg	906	27	73.0	730	2	S48813	hypothetical prote
834	27	73.0	510	1	A34160	laurate omega-hydr	907	27	73.0	742	2	T33514	hypothetical prote
835	27	73.0	510	2	T05939	cytochrome P450 m	908	27	73.0	743	2	B88551	protein T23G5.2 li
836	27	73.0	511	1	A34260	laurate omega-hydr	909	27	73.0	743	2	E71432	hypothetical prote
837	27	73.0	511	1	B34160	cytochrome P450 4A	910	27	73.0	748	2	T49633	glucan 1,4-alpha-g
838	27	73.0	511	1	B40164	cytochrome P450 4B	911	27	73.0	777	2	A35966	chemotaxis protein
839	27	73.0	511	1	O4HUB1	cytochrome P450 4B	912	27	73.0	794	2	C87437	penicillin-binding
840	27	73.0	511	2	S66472	cytochrome P450 4B	913	27	73.0	834	2	S66258	glucosidase I - hu
841	27	73.0	511	2	T00864	cytochrome P450 ho	914	27	73.0	838	2	I45557	eyeless, long form
842	27	73.0	512	2	T00605	probable cytochrom	915	27	73.0	859	2	T35785	probable beta-gluc
843	27	73.0	513	2	A12555	hypothetical prote	916	27	73.0	864	2	H85335	hypothetical prote
844	27	73.0	514	2	T08014	cytochrome P450 cy	917	27	73.0	864	2	T04518	hypothetical prote
845	27	73.0	514	2	D84584	probable protein p	918	27	73.0	873	2	A47283	calphotin - fruit
846	27	73.0	515	2	I49154	calcitonin recepto	919	27	73.0	896	2	T00389	hypothetical prote
847	27	73.0	515	2	I60800	calcitonin recepto	920	27	73.0	903	2	T09143	alpha-glucosidase
848	27	73.0	516	2	T00514	cytochrome P450 ho	921	27	73.0	913	2	S61580	probable membrane
849	27	73.0	516	2	T09364	cytochrome P450 ho	922	27	73.0	929	2	JC6124	diacylglycerol kin
850	27	73.0	519	2	I53015	fatty acid omega-h	923	27	73.0	972	2	T50400	origin recognition
851	27	73.0	519	2	JX0331	laurate omega-hydr	924	27	73.0	982	1	GNLJH2	pol polyprotein -
852	27	73.0	520	1	A46661	leukotriene B4 one	925	27	73.0	982	2	T19171	hypothetical prote
853	27	73.0	520	2	S45702	leukotriene-B4 20-	926	27	73.0	992	2	T46337	hypothetical prote
854	27	73.0	523	2	C84753	hypothetical prote	927	27	73.0	1004	2	A48821	wnt-5 protein - fr
855	27	73.0	526	2	T13687	hypothetical prote	928	27	73.0	1006	2	T00050	hypothetical prote
856	27	73.0	527	2	A84645	probable cytochrom	929	27	73.0	1006	2	G86292	hypothetical prote
857	27	73.0	531	1	SYNCYT	tyrosine-tRNA liga	930	27	73.0	1014	2	S48235	CIN1 protein - yea
858	27	73.0	533	1	S56652	calcium-dependent	931	27	73.0	1016	2	D86308	translation initia
859	27	73.0	536	2	S33569	protein-tyrosine k	932	27	73.0	1025	1	A43526	complement C3d/Eps
860	27	73.0	547	2	T03262	cytochrome P450 -	933	27	73.0	1033	2	T37715	actin-interacting
861	27	73.0	548	2	G75364	DNA mismatch repai	934	27	73.0	1036	2	D84741	probable cellulose
862	27	73.0	552	2	A96756	hypothetical prote	935	27	73.0	1042	1	GBECE	beta-galactosidase
863	27	73.0	570	2	T27407	hypothetical prote	936	27	73.0	1042	2	E85968	evolved beta-D-gal
864	27	73.0	576	2	T38293	hypothetical serin	937	27	73.0	1042	2	F91123	evolved beta-D-gal
865	27	73.0	577	2	H86414	F2H24.8 protein -	938	27	73.0	1046	1	GNWEC	genome polyprotein
866	27	73.0	578	2	E85048	hypothetical prote	939	27	73.0	1048	2	A70592	hypothetical prote
867	27	73.0	581	2	E91165	gamma-glutamyltran	940	27	73.0	1061	1	GNLJG4	HIV-1 retropepsin
868	27	73.0	581	2	E86011	protein F2D10.4 li	941	27	73.0	1066	2	T10108	vinculin - mouse
869	27	73.0	582	2	G86338	hypothetical prote	942	27	73.0	1086	2	T33893	hypothetical prote
870	27	73.0	582	2	A87316	hypothetical prote	943	27	73.0	1089	2	C70522	probable mmp18 pro
871	27	73.0	583	2	A70729	hypothetical prote	944	27	73.0	1102	2	T28666	protein kinase C-r
872	27	73.0	584	2	H87304	hypothetical prote	945	27	73.0	1107	1	JQ1658	genome polyprotein
873	27	73.0	589	2	F86549	CT365 hypothetical	946	27	73.0	1134	1	A35955	meta-vinculin - hu
874	27	73.0	589	2	D72074	conserved hypotet	947	27	73.0	1135	1	A29997	meta-vinculin - ch
875	27	73.0	589	2	T34596	probable membrane	948	27	73.0	1151	2	G96805	hypothetical prote
876	27	73.0	591	2	I65981	fatty acid omega-h	949	27	73.0	1167	2	T34020	zinc finger protei
877	27	73.0	598	2	T02795	probable membrane	950	27	73.0	1186	2	T51793	hypothetical prote
878	27	73.0	599	2	JH0259	prostaglandin-endo	951	27	73.0	1186	2	T33754	O/E-1-associated z
879	27	73.0	601	2	T49745	hypothetical prote	952	27	73.0	1199	2	A41939	G protein-coupled
880	27	73.0	605	2	H87365	sensor histidine k	953	27	73.0	1201	2	T29329	hypothetical prote
881	27	73.0	608	2	S32823	somatotropin recep	954	27	73.0	1202	2	P00440	polyprotein - barl
882	27	73.0	613	2	S27770	hypothetical prote	955	27	73.0	1205	2	T27053	hypothetical prote
883	27	73.0	616	2	JQ1441	hypothetical 67K p	956	27	73.0	1259	2	A43425	Bravo/Nr-CAM cell
884	27	73.0	616	2	I38155	DNA-binding regula	957	27	73.0	1262	2	T13353	protein stn-B - fr
885	27	73.0	632	2	E69407	NADH oxidase (nox)	958	27	73.0	1268	1	A39640	neural cell adhesi
886	27	73.0	632	2	G69306	NADH oxidase (nox)	959	27	73.0	1271	2	T24008	hypothetical prote
887	27	73.0	634	2	F82623	potassium uptake p	960	27	73.0	1280	2	T00365	hypothetical prote
888	27	73.0	637	2	A36427	lamin B receptor -	961	27	73.0	1284	2	T40879	probable helicase
889	27	73.0	640	2	S67656	hypothetical prote	962	27	73.0	1299	2	AH2090	two-component hydr
890	27	73.0	641	2	JC7331	gamma-glutamyltran	963	27	73.0	1361	2	T30884	neural specific DN
891	27	73.0	650	2	T04830	probable serine/th	964	27	73.0	1386	2	T49316	profilaggrin relat
892	27	73.0	671	2	F82960	hypothetical prote	965	27	73.0	1402	2	I46707	translation initia
893	27	73.0	674	2	S74506	ribonuclease E - S	966	27	73.0	1429	2	T13720	gene expanded prot
894	27	73.0	676	2	A45900	complement C3b rec	967	27	73.0	1443	2	S05979	steroid hormone re
895	27	73.0	678	2	B70913	probable penicilli	968	27	73.0	1445	2	A59437	KIAA1204 protein l
896	27	73.0	679	2	B75262	conserved hypotet	969	27	73.0	1494	2	T14355	protein-tyrosine-p
897	27	73.0	681	2	T17342	hypothetical prote	970	27	73.0	1520	2	T00273	hypothetical prote
898	27	73.0	695	2	S51433	MDL1 protein - yea	971	27	73.0	1522	2	T39371	transcription regu
899	27	73.0	696	2	A41344	lutropin-choriogon	972	27	73.0	1532	2	A26039	IGA-specific metal
900	27	73.0	699	1	QRHUUT	lutropin-choriogon	973	27	73.0	1561	2	S61314	IGA-specific metal
901	27	73.0	699	2	D82798	phage-related term	974	27	73.0	1603	2	S17983	gene posterior sex
902	27	73.0	702	2	JC7890	fructan beta-fruct	975	27	73.0	1693	2	AC3240	helicase, SNF2 fam
903	27	73.0	704	2	JC7932	phosphorylated car	976	27	73.0	1759	2	T18868	myoblast city prot
904	27	73.0	708	2	F87245	penicillin-binding	977	27	73.0	1763	2	S16365	collagen alpha 2(I
905	27	73.0	718	2	A13420	penicillin-binding	978	27	73.0	1773	2	A81937	IGA-specific metal

979 27 73.0 1815 2 C81169 IGA-specific metal
 980 27 73.0 1842 2 T43409 probable fatty-acid
 981 27 73.0 1842 2 T38781 fatty acid synthase
 982 27 73.0 1908 2 A86311 protein F113.14 li
 983 27 73.0 1941 2 T23979 hypothetical prote
 984 27 73.0 1943 2 T23986 hypothetical prote
 985 27 73.0 1952 2 T48814 hypothetical prote
 986 27 73.0 2175 1 S03170 homeotic protein c
 987 27 73.0 2301 2 T02323 nodulin-like prote
 988 27 73.0 2424 2 T46480 calcium channel B1
 989 27 73.0 2468 2 A83412 hypothetical prote
 990 27 73.0 2523 2 F70846 probable PPG prote
 991 27 73.0 2672 2 A48126 translation activa
 992 27 73.0 2824 2 T22759 hypothetical prote
 993 27 73.0 2897 2 B48666 cell proliferation
 994 27 73.0 3122 2 T17202 DNA-directed DNA p
 995 27 73.0 3256 2 A48666 cell proliferation
 996 27 73.0 3386 1 GNWVDF genome polyprotein
 997 27 73.0 3511 2 A59295 unconventional myo
 998 27 73.0 3716 2 E70969 probable PPG prote
 999 27 73.0 5175 2 T20992 hypothetical prote
 1000 27 73.0 5198 2 T43290 hemikentin precurs

ALIGNMENTS

RESULT 1
 B49502 protein-tyrosine-phosphatase (EC 3.1.3.48) receptor type 4E, splice form B precursor - F
 C:Species: Drosophila melanogaster
 C:Date: 10-Nov-1995 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
 C:Accession: B49502
 R:Oon, S.H.; Hong, A.; Yang, X.; Chia, W.
 J. Biol. Chem. 268, 23964-23971, 1993
 A>Title: Alternative splicing in a novel tyrosine phosphatase gene (DTP4E) of Drosophila
 A:Reference number: A49502; MUID:94043220; PMID:8226938
 A:Accession: B49502

A:Molecule type: mRNA
 A:Residues: 1-1615 <CON>
 A:Cross-references: UNIPROT:Q9W4F5; UNIPARC:UPI0000177053; GB:L20894
 C:Genetics:
 A:Gene: FlyBase:Ptp4E
 A:Cross-references: FlyBase:FBgn0004368
 A:Introns: 1605/3

C:Superfamily: protein-tyrosine-phosphatase, receptor type 4E; fibronectin type III repe
 C:Keywords: alternative splicing; phosphoprotein; phosphoric monoester hydrolase; recept
 F:1254-1270/Domain: transmembrane #status predicted <TMN>
 F:1271-1615/Domain: intracellular #status predicted <INT>
 F:1353-1573/Domain: protein-tyrosine-phosphatase homology <PTPI>
 F:1525/Active site: Cys (phosphocysteine intermediate) #status predicted
 F:1531/Binding site: substrate phosphate (Arg) #status predicted

Query Match 100.0%; Score 37; DB 2; Length 1615;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02; Indels 0; Gaps 0;
 Matches 7; Conservative 0; Mismatches 0;

QY 1 ILAPPVP 7
 |||||
 Db 1101 ILAPPVP 1107

RESULT 2
 A49502 protein-tyrosine-phosphatase (EC 3.1.3.48), receptor type 4E, splice form A precursor -
 C:Species: Drosophila melanogaster
 C:Date: 10-Nov-1995 #sequence_revision 10-Nov-1995 #text_change 31-Dec-2004
 C:Accession: A49502
 R:Oon, S.H.; Hong, A.; Yang, X.; Chia, W.
 J. Biol. Chem. 268, 23964-23971, 1993
 A>Title: Alternative splicing in a novel tyrosine phosphatase gene (DTP4E) of Drosophila
 A:Reference number: A49502; MUID:94043220; PMID:8226938
 A:Accession: A49502

A:Molecule type: mRNA
 A:Residues: 1-1767 <CON>
 A:Cross-references: UNIPROT:Q9W4F5; UNIPARC:UPI0000177052; GB:L20894
 A:Note: authors translated the codon ATA for residue 1715 as Leu
 C:Genetics:
 A:Gene: ptp4E

A:Cross-references: FlyBase:FBgn0004368
 C:Superfamily: fibronectin type III repeat homology; protein-tyrosine-phosphatase homolo-
 C:Keywords: alternative splicing; phosphoprotein; phosphoric monoester hydrolase; recept
 F:1254-1270/Domain: transmembrane #status predicted <TMN>
 F:1271-1767/Domain: intracellular #status predicted <INT>
 F:1353-1573/Domain: protein-tyrosine-phosphatase homology <PTPI>
 F:1525/Active site: Cys (phosphocysteine intermediate) #status predicted
 F:1531/Binding site: substrate phosphate (Arg) #status predicted

Query Match 100.0%; Score 37; DB 2; Length 1767;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAPPVP 7
 |||||
 Db 1101 ILAPPVP 1107

RESULT 3
 F70595 hypothetical protein RV3213c - Mycobacterium tuberculosis (strain H37RV)
 C:Species: Mycobacterium tuberculosis
 C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
 C:Accession: F70595
 R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.
 ; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
 Rajandram, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
 Nature 393, 537-544, 1998

A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A>Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
 A:Reference number: A70500; MUID:98295987; PMID:9634230
 A:Accession: F70595
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-266 <COL>
 A:Cross-references: UNIPROT:O05853; UNIPARC:UPI00000D4F2A; GB:Z95120; GB:AL123456; NID:9
 A:Experimental source: strain H37RV
 C:Genetics:
 A:Gene: RV3213c
 C:Superfamily: regulatory protein spoJ

Query Match 94.6%; Score 35; DB 2; Length 266;
 Best Local Similarity 71.4%; Pred. No. 38;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAPPVP 7
 :|||:
 Db 209 VLAPPVP 215

RESULT 4
 T24591 hypothetical protein T06E4.7 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
 C:Accession: T24591
 R:Lloyd, C.
 submitted to the EMBL Data Library, April 1996
 A:Reference number: Z19910
 A:Accession: T24591
 A:Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: DNA
 A:Residues: 1-441 <WIL>
 A:Cross-references: UNIPROT:Q22261; UNIPARC:UPI0000082B31; EMBL:Z70756; PIDN:CAA94793.1;
 A:Experimental source: clone T06E4
 C:Genetics:
 A:Gene: CESP:T06E4.7

A:Map position: 5
A:Introns: 104/2; 317/2; 354/3; 404/3

Query Match 94.6%; Score 35; DB 2; Length 441;
Best Local Similarity 85.7%; Pred. No. 65;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAPPVP 7
:|||||
Db 210 LIAPPVP 216

RESULT 5
D71658
3-octaprenyl-4-hydroxybenzoate carboxy-lyase (ubiX) RP541 - Rickettsia prowazekii
C:Species: Rickettsia prowazekii
C:Date: 21-Nov-1998 #sequence_revision 21-Nov-1998 #text_change 09-Jul-2004
C:Accession: D71658
R:Andersson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Sicheritz-Ponten, T.; Alsmark, U.
Nature 396, 133-140, 1998
A:Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.
A:Reference number: A71630; MUID:99039499; PMID:9823893
A:Accession: D71658
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-189 <AND>
A:Cross-references: UNIPROT:Q9ZD09; UNIPARC:UPI0000131197; GB:AJ235269; NID
A:Experimental source: strain Madrid E
C:Genetics:
A:Gene: ubiX; RP541
C:Superfamily: dedF protein

Query Match 91.9%; Score 34; DB 2; Length 189;
Best Local Similarity 71.4%; Pred. No. 40;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAPPVP 7
:|||||
Db 146 LIAPPVP 152

RESULT 6
S75284
chemotaxis protein cheA homolog - Synechocystis sp. (strain PCC 6803)
N:Alternate names: protein sll1296
C:Species: Synechocystis sp.
A:Variety: PCC 6803
C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C:Accession: S75284
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
sp.
A:Reference number: S74322; MUID:97061201; PMID:8905231
A:Accession: S75284
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-924 <KAN>
A:Cross-references: UNIPROT:P73172; UNIPARC:UPI00000D34AC; EMBL:D30904; GB:AB001339; NID
A:Note: The nucleotide sequence was submitted to the EMBL Data Library, June 1996
C:Genetics:
A:Gene: cheA
C:Keywords: chemotaxis; phosphoprotein
F:802-914/Domain: response regulator homology <RRH>
F:851/Binding site: phosphate (Asp) (covalent) #status predicted

Query Match 91.9%; Score 34; DB 2; Length 924;
Best Local Similarity 71.4%; Pred. No. 2.2e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAPPVP 7
:|||||

Db 279 VIAPPVP 285

RESULT 7
AE1448
hypothetical protein lin0124 [imported] - Listeria innocua (strain Clip11262)
C:Species: Listeria innocua
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C:Accession: AE1448
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
Science 294, 849-852, 2001
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland
A:Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AE1448
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-113 <GLA>
A:Cross-references: UNIPROT:Q925X3; UNIPARC:UPI00001392F0; GB:AL592022; PIDN:CAC95357.1;
A:Experimental source: strain Clip11262
C:Genetics:
A:Gene: lin0124

Query Match 89.2%; Score 33; DB 2; Length 113;
Best Local Similarity 71.4%; Pred. No. 35;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAPPVP 7
:|||||
Db 36 LIAPPVP 42

RESULT 8
AE1729
hypothetical protein lin2378 [imported] - Listeria innocua (strain Clip11262)
C:Species: Listeria innocua
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C:Accession: AE1729
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland
A:Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AE1729
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-113 <GLA>
A:Cross-references: UNIPROT:Q925X3; UNIPARC:UPI00001392F0; GB:AL592022; PIDN:CAC97605.1;
A:Experimental source: strain Clip11262
C:Genetics:
A:Gene: lin2378

Query Match 89.2%; Score 33; DB 2; Length 113;
Best Local Similarity 71.4%; Pred. No. 35;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAPPVP 7
:|||||
Db 36 LIAPPVP 42

RESULT 9
AG1752
hypothetical protein lin2564 [imported] - Listeria innocua (strain Clip11262)
C:Species: Listeria innocua
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C:Accession: AG1752

R.; Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.
 Science 294, 849-852, 2001
 A; Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Møller, C.; Schlüter, T.; Simoes, N.; Tietz, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.; Title: Comparative genomics of *Listeria* species.
 A; Reference number: AB1077; MUID:21537279; PMID:11679669
 A; Accession: AG1752
 A; Status: preliminary
 A; Molecule type: DNA
 A; Residues: 1-113 <GLA>
 A; Cross-references: UNIPROT:Q925X3; UNIPARC:UPI00001392F0; GB:AL592022; PIDN:CAC97791.1;
 A; Experimental source: strain Clp11262
 C; Genetics:
 A; Gene: lin2564

Query Match 89.2%; Score 33; DB 2; Length 113;
 Best Local Similarity 71.4%; Pred. No. 35;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAPPVP 7
 : : : : :
 Db 36 LIAPPVP 42

RESULT 10
 S41002
 Hypothetical protein T05G5.2 - *Caenorhabditis elegans*
 C; Species: *Caenorhabditis elegans*
 C; Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
 C; Accession: S41002
 R; Thomas, K.
 submitted to the EMBL Data Library, October 1993
 A; Reference number: S41001
 A; Accession: S41002
 A; Status: preliminary
 A; Molecule type: DNA
 A; Residues: 1-205 <THO>
 A; Cross-references: UNIPROT:P34555; UNIPARC:UPI0000061154; EMBL:D227079; NID:G414641; PID
 C; Genetics:
 A; Introns: 21/3

Query Match 89.2%; Score 33; DB 2; Length 205;
 Best Local Similarity 100.0%; Pred. No. 66;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LAPPVP 7
 : : : : :
 Db 139 LAPPVP 144

RESULT 11
 T35712
 Integral membrane protein - *Streptomyces coelicolor*
 C; Species: *Streptomyces coelicolor*
 C; Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 09-Jul-2004
 C; Accession: T35712
 R; Murphy, L.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
 submitted to the EMBL Data Library, January 1998
 A; Reference number: Z21548
 A; Accession: T35712
 A; Status: preliminary; translated from GB/EMBL/DBJ
 A; Molecule type: DNA
 A; Residues: 1-465 <MUR>
 A; Cross-references: UNIPROT:Q54176; UNIPARC:UPI00000DABB3; EMBL:AL021411; PIDN:CAA16204.
 A; Experimental source: strain A3(2)
 C; Genetics:
 A; Gene: SC0EDB:SC7H1.17
 C; Superfamily: multidrug-efflux transporter

Query Match 89.2%; Score 33; DB 2; Length 465;
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LAPPVP 7
 : : : : :
 Db 5 LAPPVP 10

RESULT 12
 S76481
 Hypothetical protein - *Synechocystis* sp. (strain PCC 6803)
 C; Species: *Synechocystis* sp.
 A; Variety: PCC 6803
 C; Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
 C; Accession: S76481
 R; Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.; O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda, DNA Res. 3, 109-136, 1996
 A; Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp.
 A; Reference number: S74322; MUID:97061201; PMID:8905231
 A; Accession: S76481
 A; Status: preliminary
 A; Molecule type: DNA
 A; Residues: 1-538 <KAN>
 A; Cross-references: UNIPROT:P74506; UNIPARC:UPI00000C104B; EMBL:D90915; GB:AB001339; NID
 A; Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 89.2%; Score 33; DB 2; Length 538;
 Best Local Similarity 85.7%; Pred. No. 1.9e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ILAPPVP 7
 : : : : :
 Db 426 IFAPPVP 432

RESULT 13
 AF1927
 Hypothetical protein all0969 [imported] - *Nostoc* sp. (strain PCC 7120)
 C; Species: *Nostoc* sp. PCC 7120
 A; Note: *Nostoc* sp. strain PCC 7120 is a synonym of *Anabaena* sp. strain PCC 7120
 C; Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
 C; Accession: AF1927
 R; Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S. DNA Res. 8, 205-213, 2001
 A; Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium *Anabaena*
 A; Reference number: AB1807; MUID:21595285; PMID:11759840
 A; Accession: AF1927
 A; Status: preliminary
 A; Molecule type: DNA
 A; Residues: 1-543 <KUR>
 A; Cross-references: UNIPROT:Q8YV82; UNIPARC:UPI00000CDF2B; GB:BA0000019; PIDN:BA072926.1;
 A; Experimental source: strain PCC 7120
 C; Genetics:
 A; Gene: all0969

Query Match 89.2%; Score 33; DB 2; Length 543;
 Best Local Similarity 85.7%; Pred. No. 1.9e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ILAPPVP 7
 : : : : :
 Db 428 IFAPPVP 434

RESULT 14
 T13994
 envelope protein - fruit fly (*Drosophila melanogaster*) ZAM retrovirus-like element
 C; Species: *Drosophila melanogaster*
 C; Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
 C; Accession: T13994
 R; Leblanc, P.; Desset, S.; Dastugue, B.; Vaurcy, C.

```

EMBO J. 16, 7521-7531, 1997
A:Title: Invertebrate retroviruses: ZAM a new candidate in Drosophila melanogaster.
A:Reference number: Z17849; MUID:98070328; PMID:9405380
A:Accession: T13994
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-551 <LEB>
A:Cross-references: UNIPROT:Q46113; UNIPARC:UPI000007A36F; EMBL:AJ000387; NID:92791286;
C:Genetics:
A:Gene: env
A:Cross-references: FlyBase:FBgn0023131
A:Mobile element: ZAM retrovirus-like element
A:Introns: 13/2

Query Match      89.2%; Score 33; DB 2; Length 551;
Best Local Similarity 71.4%; Pred. No. 1.9e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      1 ILAPPVP 7
      |||||
Db      520 IVAPPFP 526

RESULT 15
T08877
Modin - Podospora anserina
C:Species: Podospora anserina
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
C:Accession: T08877
R:Barreau, C.; Iskandar, M.; Loubradou, G.; Levallois, V.; Begueret, J.
Genetics 149, 915-926, 1998
A:Title: The mod-A suppressor of nonallelic heterokaryon incompatibility in Podospora anserina
A:Reference number: Z16508; MUID:98278809; PMID:9611202
A:Accession: T08877
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-687 <BAR>
A:Cross-references: UNIPROT:Q59895; UNIPARC:UPI000006C81B; EMBL:AF025289; NID:G3115380;
A:Experimental source: strain het-cl het-el het-dl mod-B1
C:Genetics:
A:Gene: mod-A
A:Introns: 358/1
C:Superfamily: Podospora anserina Modin

Query Match      89.2%; Score 33; DB 2; Length 687;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 LAPPVP 7
      |||||
Db      599 LAPPVP 604

RESULT 16
F87708
cell division protein FtsK, probable [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C:Accession: F87708
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: F87708
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-819 <STO>
A:Cross-references: UNIPROT:Q9A262; UNIPARC:UPI00000C7B8B; GB:AE005673; NID:gl3425470; F
C:Genetics:
A:Gene: CC3704

```

```

Query Match      89.2%; Score 33; DB 2; Length 819;
Best Local Similarity 85.7%; Pred. No. 2.9e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 ILAPPVP 7
      |||||
Db      811 ILAPPFP 817

RESULT 17
A41401
Mineralocorticoid receptor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 03-Apr-1992 #sequence_revision 03-Apr-1992 #text_change 05-Oct-2004
C:Accession: A41401
R:Patel, P.D.; Sherman, T.G.; Goldman, D.J.; Watson, S.J.
Mol. Endocrinol. 3, 1877-1885, 1989
A:Title: Molecular cloning of a mineralocorticoid (type I) receptor complementary DNA fr
A:Reference number: A41401; MUID:90114194; PMID:2558305
A:Accession: A41401
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-981 <PAT>
A:Cross-references: UNIPROT:P22199; UNIPARC:UPI000012EDE6; GB:M36074; NID:G205340; PIDN.
C:Superfamily: Mineralocorticoid receptor (MR); erba transforming protein homology
C:Keywords: DNA binding; transcription regulation; zinc finger
F:602-877/Domain: erba transforming protein homology <ERBA>
F:604-624/Region: zinc finger
F:640-664/Region: zinc finger

Query Match      89.2%; Score 33; DB 2; Length 981;
Best Local Similarity 85.7%; Pred. No. 3.6e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 ILAPPVP 7
      |||||
Db      473 ILGPPVP 479

RESULT 18
A29513
Mineralocorticoid receptor - human
N:Alternate names: aldosterone receptor
C:Species: Homo sapiens (man)
C:Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 05-Oct-2004
C:Accession: A29513
R:Arriza, J.L.; Weinberger, C.; Cerelli, G.; Glaser, T.M.; Handelin, B.L.; Houseman, D.E.
Science 237, 268-275, 1987
A:Title: Cloning of human mineralocorticoid receptor complementary DNA: structural and f
A:Reference number: A29513; MUID:87263386; PMID:3037703
A:Accession: A29513
A:Molecule type: mRNA
A:Residues: 1-984 <ARR>
A:Cross-references: UNIPROT:P08235; UNIPARC:UPI000012EDE5; GB:M16801; NID:gl87460; PIDN:
C:Genetics:
A:Gene: GDB:MLR
A:Cross-references: GDB:120188; OMIM:264350
A:Map position: 4q31-q31
C:Superfamily: Mineralocorticoid receptor (MR); erba transforming protein homology
C:Keywords: DNA binding; transcription regulation; zinc finger
F:601-880/Domain: erba transforming protein homology <ERBA>
F:603-623/Region: zinc finger
F:639-663/Region: zinc finger

Query Match      89.2%; Score 33; DB 2; Length 984;
Best Local Similarity 85.7%; Pred. No. 3.6e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 ILAPPVP 7
      |||||
Db      472 ILGPPVP 478

```

RESULT 19

T21913
Hypothetical protein F44G4.8 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T21913; T22211
R:Thomas, K.
submitted to the EMBL Data Library, September 1995
A:Reference number: Z19486
A:Accession: T21913
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1367 <W1>
A:Cross-references: UNIPROT:Q20120; UNIPARC:UPI0000080BD3; EMBL:Z54218; PIDN:CAA90958.1;
A:Experimental source: clone F37B12
R:Sims, M.
submitted to the EMBL Data Library, June 1995
A:Reference number: Z19530
A:Accession: T22211
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1367 <W12>
A:Cross-references: UNIPARC:UPI0000080BD3; EMBL:Z49910; PIDN:CAA90125.1; GSPDB:GN000020;
A:Experimental source: clone F44G4
C:Genetics:
A:Gene: CESP:F44G4.8
A:Map position: 2
A:Intron: 76/2; 109/2; 163/3; 229/1; 276/3; 329/3; 381/2; 430/3; 648/1; 780/2; 805/3; 8

Query Match 89.2%; Score 33; DB 2; Length 1367;
Best Local Similarity 71.4%; Pred. No. 5.1e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILAPPVP 7
:::|||||
Db 740 LMAPPVP 746

RESULT 20

I38346
elastic titin - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 09-Jul-2004
C:Accession: I38346
R:Labat, S.; Kolmerer, B.
Science 270, 293-296, 1995
A:Title: Titins: Giant proteins in charge of muscle ultrastructure and elasticity.
A:Reference number: A57430; MUID:96026330; PMID:7569978
A:Accession: I38346
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-7962 <RES>
A:Cross-references: UNIPROT:Q10465; UNIPARC:UPI000011010C; EMBL:X90569; NID:g1017426; PI
C:Genetics:
A:Gene: GDB:TTN
A:Map position: 2q31-2q31

Query Match 89.2%; Score 33; DB 2; Length 7962;
Best Local Similarity 71.4%; Pred. No. 3.3e+03;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILAPPVP 7
:::|||||
Db 5625 VVAPPVP 5631

RESULT 21

A37800
Hypothetical protein ubix [imported] - Rickettsia conorii (strain Malish 7)
C:Species: Rickettsia conorii
C>Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C:Accession: A97800

R:Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.; Ro
Science 293, 2093-2098, 2001
A:Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.
A:Reference number: A97700; MUID:21442074; PMID:11557893

A:Accession: A97800
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-191 <KUR>
A:Cross-references: UNIPROT:Q92HH0; UNIPARC:UPI000000CBEE1; GB:AE006914; PIDN:AAL03339.1;
C:Genetics:
A:Gene: ubix
C:Superfamily: dedF protein

Query Match 86.5%; Score 32; DB 2; Length 191;
Best Local Similarity 71.4%; Pred. No. 93;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILAPPVP 7
:::|||||
Db 146 IISPPVP 152

RESULT 22

H72585
Hypothetical protein APE1157 - Aeropyrum pernix (strain K1)
C:Species: Aeropyrum pernix
C>Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C:Accession: H72585
R:Kawarayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K
DNA Res. 6, 83-101, 1999
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr
A:Reference number: A72450; MUID:99310339; PMID:10382966
A:Accession: H72585
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-210 <KAW>
A:Cross-references: UNIPROT:Q9YCV5; UNIPARC:UPI000005DE2B; DBJ:AP000061; NID:G5104821;
A:Experimental source: strain K1
C:Genetics:
A:Gene: APE1157

Query Match 86.5%; Score 32; DB 2; Length 210;
Best Local Similarity 83.3%; Pred. No. 1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LAPPVP 7
:::|||||
Db 95 LAPPVP 100

RESULT 23

T10399
Hypothetical protein 130 - Orgyia pseudotsugata nuclear polyhedrosis virus
C:Species: Orgyia pseudotsugata nuclear polyhedrosis virus, OpNPV
C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
C:Accession: T10399
R:Ahrens, C.A.; Russell, R.R.; Funk, C.J.; Evans, J.; Harwood, S.; Rohrmann, G.F.
Virology 229, 381-399, 1997
A:Title: The sequence of the Orgyia pseudotsugata multinucleocapsid nuclear polyhedrosis
A:Reference number: Z17011; MUID:97271300; PMID:9126251
A:Accession: T10399
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-228 <AHR>
A:Cross-references: UNIPARC:UPI000006186E; EMBL:U75930; NID:g2934903; PIDN:AAC59129.1; P
C:Superfamily: OpNPV hypothetical protein 130

Query Match 86.5%; Score 32; DB 2; Length 228;
Best Local Similarity 83.3%; Pred. No. 1.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LAPPVP 7

```

Db          218 LAPPVP 223
|||||:|
Query Match      86.5%; Score 32; DB 2; Length 228;
Best Local Similarity 83.3%; Pred. NO. 1.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

RESULT 24
D30857
hypothetical protein 3 - Orgyia pseudotsugata multicapsid nuclear polyhedrosis virus
C:Species: Orgyia pseudotsugata multicapsid nuclear polyhedrosis virus, OpMNPV
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 09-Jul-2004
C:Accession: D30857; J70431
R:Gombart, A.F.; Blissard, G.W.; Rohrmann, G.F.
submitted to GenBank, February 1989
A:Reference number: A30857
A:Accession: D30857
A:Molecule type: DNA
A:Residues: 1-228 <OM>
A:Cross-references: UNIPROT:P24080; UNIPARC:UPI000006186E; GB:D13929; GB:D00508; NID:g22
C:Genetics:
A:Map position: 84.45-87.70
C:Superfamily: OpMNPV hypothetical protein 130

Query Match      86.5%; Score 32; DB 2; Length 228;
Best Local Similarity 83.3%; Pred. NO. 1.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

RESULT 25
E86433
protein T17H7.3 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cross)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:Accession: E86433
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, J.R.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: E86433
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-481 <STO>
A:Cross-references: UNIPROT:Q9SY17; UNIPARC:UPI00000A0CD1; GB:AE005172; NID:g4926818; PI
C:Genetics:
A:Gene: t17H7.3
A:Map position: 1

Query Match      86.5%; Score 32; DB 2; Length 481;
Best Local Similarity 71.4%; Pred. NO. 2.5e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY          1 ILAPPVP 7
:||||:|
Db          173 LLAPPLP 179

RESULT 26
T24393
hypothetical protein T03E6.8 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T24393
R:Lloyd, C.
submitted to the EMBL Data Library, March 1997

```

```

A:Reference number: Z19885
A:Accession: T24393
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-581 <WIL>
A:Cross-references: UNIPROT:O45740; UNIPARC:UPI000016427C; EMBL:Z92812; PIDN:CAB07281.1;
A:Experimental source: clone T03E6
C:Genetics:
A:Gene: CESP:T03E6.8
A:Map position: 5
A:Introns: 60/2; 113/1; 182/3; 222/3; 249/1; 305/2; 354/1; 540/1; 566/2

Query Match      86.5%; Score 32; DB 2; Length 581;
Best Local Similarity 85.7%; Pred. NO. 3.1e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY          1 ILAPPVP 7
:||||:|
Db          320 ILEPPVP 326

RESULT 27
C86152
T7123.2 protein - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cross)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:Accession: C86152
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: C86152
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-642 <STO>
A:Cross-references: UNIPROT:O23673; UNIPARC:UPI000009D3C0; GB:AE005172; NID:g2317902; PI
C:Genetics:
A:Map position: 1

Query Match      86.5%; Score 32; DB 2; Length 642;
Best Local Similarity 57.1%; Pred. NO. 3.4e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY          1 ILAPPVP 7
:||||:|
Db          186 LIAPPVP 192

RESULT 28
A96807
hypothetical protein T32E8.5 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cross)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:Accession: A96807
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: A96807

```

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-720 <STO>
A:Cross-references: UNIPROT:O9CA22; UNIPARC:UPI000009F53F; GB:AE005173; NID:g6437532; PI
C:Genetics:
A:Gene: T32B8.5
A:Map position: 1

Query Match 86.5%; Score 32; DB 2; Length 720;
Best Local Similarity 83.3%; Pred. No. 3.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LAPPVP 7
DB 636 LAPPVP 641

RESULT 29
T01619
hypothetical protein At2g18910 [imported] - Arabidopsis thaliana
N:Alternate names: hypothetical protein F19F24.11
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 19-Feb-1999 #sequence_revision 19-Feb-1999 #text_change 09-Jul-2004
A:Accession: T01619; B84570
R:Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul
submitted to the EMBL Data Library, April 1998
A:Description: Arabidopsis thaliana chromosome II BAC F19F24 genomic sequence.
A:Reference number: Z14153
A:Accession: T01619
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-753 <ROU>
A:Cross-references: UNIPROT:O64621; UNIPARC:UPI000017AF77; EMBL:AC003673; NID:g3004543;
A:Experimental source: cultivar Columbia
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Unayam, L.; Tallon, L.;
euse, D.; Niernan, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: B84570
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-753 <STO>
A:Cross-references: UNIPARC:UPI000017AF77; GB:AE002093; NID:g3004565; PIDN:AAC09038.1; G
C:Genetics:
A:Gene: At2g18910; F19F24.11
A:Map position: 2
A:Introns: 78/2; 606/1; 705/1; 737/3

Query Match 86.5%; Score 32; DB 2; Length 753;
Best Local Similarity 83.3%; Pred. No. 4.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LAPPVP 7
DB 388 LAPPVP 393

RESULT 30
T00377
K1AA0642 protein - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 09-Jul-2004
A:Accession: T00377
R:Ishikawa, K.; Nagase, T.; Suyama, M.; Miyajima, N.; Tanaka, A.; Kotani, H.; Nomura, N.
DNA Res. 5, 169-176, 1998
A:Title: Prediction of the coding sequences of unidentified human genes. X. The complete
A:Reference number: Z14142; MUID:98403880; PMID:9734811
A:Accession: T00377
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1069 <ISH>

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-720 <STO>
A:Cross-references: UNIPROT:O75137; UNIPARC:UPI000017C289; EMBL:AB014542; NID:g3327097;
C:Genetics:
A:Gene: KIAA0642

Query Match 86.5%; Score 32; DB 2; Length 1069;
Best Local Similarity 85.7%; Pred. No. 5.9e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ILAPPVP 7
DB 220 ILAPPVP 226

RESULT 31
D89451
protein T04G9.1 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
A:Accession: D89451
R:Anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog
A:Reference number: A75000; MUID:99069613; PMID:9851916
A:Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_ele
A:Note: published errata appeared in Science 283, 35, 1999, Science 283, 2103, 1999, and
A:Accession: D89451
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1610 <STO>
A:Cross-references: UNIPROT:Q22173; UNIPARC:UPI000007A764; GB:chr_X; PIDN:AAA82463.1; PI
C:Genetics:
A:Gene: T04G9.1
A:Map position: X

Query Match 86.5%; Score 32; DB 2; Length 1610;
Best Local Similarity 85.7%; Pred. No. 9.2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ILAPPVP 7
DB 43 ILAPPVP 49

RESULT 32
A75613
hypothetical protein DRA0166 - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
A:Accession: A75613
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vamathevan, J.G.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896; PMID:10567266
A:Accession: A75613
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1626 <WHI>
A:Cross-references: UNIPROT:Q9RYV7; UNIPARC:UPI00000C164F; GB:AE001862; GB:AE001825; NID
A:Experimental source: strain R1
C:Genetics:
A:Gene: DRA0166
A:Map position: 2
C:Superfamily: Deinococcus radiodurans hypothetical protein DRA0166

Query Match 86.5%; Score 32; DB 2; Length 1626;
Best Local Similarity 71.4%; Pred. No. 9.3e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ILAPPVP 7
DB 752 VLAPPVP 758

RESULT 33
T00390
K1AA0614 protein - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 09-Jul-2004
A:Accession: T00390; T17254
R:Ishikawa, K.; Nagase, T.; Suyama, M.; Miyajima, N.; Tanaka, A.; Kotani, H.; Nomura, N.
DNA Res. 5, 169-176, 1998
A>Title: Prediction of the coding sequences of unidentified human genes. X. The complete
A:Reference number: Z14142; MUID:98403880; PMID:9734811
A:Accession: T00390
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1630 <ISH>
A:CROSS-references: UNIPROT:Q9UFT6; UNIPARC:UPI000003B446; EMBL:AB014514; NID:G3327041;
R:Koehrer, K.; Beyer, A.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, September 1999
A:Reference number: Z18722
A:Accession: T17254
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1515-1630 <KOR>
A:CROSS-references: UNIPARC:UPI0000073F07; EMBL:AL117469
A:Experimental source: adult uterus; clone DKFZps8601022
C:Genetics:
A>Note: KIAA0614; DKFZps8601022.1

Query Match 86.5%; Score 32; DB 2; Length 1630;
Best Local Similarity 71.4%; Pred. No. 9.3e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAPVP 7
:||||:
DB 122 LLAAPLP 128

RESULT 34
T24089
hypothetical protein R09E10.7 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T24089
R:Matthews, L.
submitted to the EMBL Data Library, March 1996
A:Reference number: Z19839
A:Accession: T24089
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1791 <WIL>
A:CROSS-references: UNIPROT:Q21875; UNIPARC:UPI000007D7B4; EMBL:Z70287; PIDN:CAA94301.1;
A:Experimental source: clone R09E10
C:Genetics:
A:Gene: CESP-R09E10.7
A:Map position: 4
A:Introns: 78/2; 94/3; 250/2; 350/2; 674/3; 1079/2; 1155/2; 1450/3; 1593/3; 1642/2; 1690

Query Match 86.5%; Score 32; DB 2; Length 1791;
Best Local Similarity 83.3%; Pred. No. 1e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LAPPVP 7
:||||:
DB 592 LAPPVP 597

RESULT 35
A59266
unconventional myosin-15 - human
C:Species: Homo sapiens (man)
C>Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 09-Jul-2004
C:Accession: A59266

R:Liang, Y.; Wang, A.; Belyantseva, I.A.; Anderson, D.W.; Probst, F.J.; Barber, T.D.; Mi
an, T.B.; Fridell, R.A.
Genomics 61, 243-258, 1999
A>Title: Characterization of the human and mouse unconventional myosin XV genes responsi
A:Reference number: A59266; MUID:20021762; PMID:10552926
A:Accession: A59266
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-3530 <LIA>
A:CROSS-references: UNIPROT:Q9UKN7; UNIPARC:UPI000012FBABF; GB:AF144094; NID:G6224682; PI
F1225-1887/Domain: myosin motor domain homology <MNO>

Query Match 86.5%; Score 32; DB 2; Length 3530;
Best Local Similarity 71.4%; Pred. No. 2.1e+03;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAPVP 7
:||||:
DB 660 LLSPPVP 666

RESULT 36
H87656

3-octaprenyl-4-hydroxybenzoate carboxy-lyase [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C>Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C:Accession: H87656
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Emolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A>Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: H87656
A:Molecule type: DNA
A>Status: preliminary
A:Residues: 1-162 <STO>
A:CROSS-references: UNIPROT:Q9A3B4; UNIPARC:UPI00000C7A0D; GB:AE005673; NID:G13424982; P
C:Genetics:
A:Gene: CC3290

Query Match 83.8%; Score 31; DB 2; Length 162;
Best Local Similarity 57.1%; Pred. No. 1.2e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAPVP 7
:||||:
DB 109 VIAPPLP 115

RESULT 37
H89164

protein T06B4.8 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C:Accession: H89164
R:Anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998

A>Title: Genome sequence of the nematode C. elegans: a platform for investigating biology
A:Reference number: A75000; MUID:99069613; PMID:9851916
A>Note: see websites genome.wustl.edu/gsc/C_elegans/ and www_sanger.ac.uk/Projects/C_ele
A>Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A:Accession: H89164
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-198 <STO>
A:CROSS-references: UNIPROT:Q22262; UNIPARC:UPI000017A685; GB:chr_V; PIDN:CAA94794.1; PI
C:Genetics:
A:Gene: T06B4.8
A:Map position: 5

Query Match 83.8%; Score 31; DB 2; Length 198;
Best Local Similarity 71.4%; Pred. No. 1.5e+02;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ILAPPVP 7
:|||||

Db 103 LLAPPAP 109

RESULT 38

T24595

Hypothetical protein T06E4.8 - *Caenorhabditis elegans*C:Species: *Caenorhabditis elegans*

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C:Accession: T24595

R:Lloyd, C.

submitted to the EMBL Data Library, April 1996

A:Reference number: Z19910

A:Accession: T24595

A>Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-199 <WIL>

A:Cross-references: UNIPROT:Q22262; UNIPARC:UPI0000081907; EMBL:Z70756; PIDN:CAA94794.2;

A:Experimental source: clone T06E4

C:Genetics:

A:Gene: CESP:T06E4.8

A:Map position: 5

A:Introns: 21/2; 198/2

Query Match

83.8%; Score 31; DB 2; Length 199;

Best Local Similarity 71.4%; Pred. No. 1.5e+02;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ILAPPVP 7

:|||||

Db 103 LLAPPAP 109

RESULT 39

T51884

Hypothetical protein DKFZp547K054.1 - human (fragment)

C:Species: *Homo sapiens* (man)

C:Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 09-Jul-2004

C:Accession: T51884

R:Bloeker, H.; Bocher, M.; Brandt, P.; Mewes, H.W.; Weil, B.; Wiemann, S.

submitted to the Protein Sequence Database, July 2000

A:Reference number: Z25853

A:Accession: T51884

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-205 <AAA>

A:Cross-references: UNIPROT:Q9NPT3; UNIPARC:UPI000006E1B0; EMBL:AL390175

A:Experimental source: fetal brain; clone DKFZp547K054

C:Genetics:

A>Note: DKFZp547K054.1

Query Match

83.8%; Score 31; DB 2; Length 205;

Best Local Similarity 71.4%; Pred. No. 1.5e+02;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ILAPPVP 7

:|||||

Db 55 LLCPPVP 61

RESULT 40

G95858

probable 3-octaprenyl-4-hydroxybenzoate carboxy-lyase protein (EC 4.1.1.-) [imported] -

C:Species: *Sinorhizobium meliloti*

C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004

C:Accession: G95858

R:Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan

Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001

A>Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing endo

A:Reference number: A95842; MUID:21396508; PMID:11481431

A:Accession: G95858
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-210 <KUR>
A:Cross-references: UNIPROT:Q92X27; UNIPARC:UPI00000CB414; GB:AL591985; PIDN:CAC48535.1;
A:Experimental source: strain 1021, megaplasmid pSymB
R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
Pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001

A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
heault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.;
A>Title: The composite genome of the legume symbiont *Sinorhizobium meliloti*.

A:Reference number: A96039; MUID:21368234; PMID:11474104

A:Contents: annotation

C:Genetics:

A:Gene: SMB20135

A:Genome: plasmid

C:Superfamily: dedF protein

C:Keywords: carbon-carbon lyase; carboxy-lyase

Query Match

83.8%; Score 31; DB 2; Length 210;

Best Local Similarity 71.4%; Pred. No. 1.6e+02;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ILAPPVP 7

:|||||

Db 150 IICPPVP 156

RESULT 41

T24594

Hypothetical protein T06E4.9 - *Caenorhabditis elegans*C:Species: *Caenorhabditis elegans*

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C:Accession: T24594

R:Lloyd, C.

submitted to the EMBL Data Library, April 1996

A:Reference number: Z19910

A:Accession: T24594

A>Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-212 <WIL>

A:Cross-references: UNIPROT:Q22263; UNIPARC:UPI00000751FF; EMBL:Z70756; PIDN:CAA94795.2;

A:Experimental source: clone T06E4

C:Genetics:

A:Gene: CESP:T06E4.9

A:Map position: 5

A:Introns: 34/2; 211/2

Query Match

83.8%; Score 31; DB 2; Length 212;

Best Local Similarity 71.4%; Pred. No. 1.6e+02;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ILAPPVP 7

:|||||

Db 116 LLAPPAP 122

RESULT 42

G99164

protein T06E4.9 [imported] - *Caenorhabditis elegans*C:Species: *Caenorhabditis elegans*

C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004

C:Accession: G99164

R:anonymous, The C. elegans Sequencing Consortium.

Science 282, 2012-2018, 1998

A>Title: Genome sequence of the nematode *C. elegans*: a platform for investigating biolog

A:Reference number: A75000; MUID:99069613; PMID:9851916

A>Note: see websites genome.wustl.edu/gsc/c_elegans/ and www.sanger.ac.uk/Projects/C_ele

A>Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and

A:Accession: G99164

A>Status: preliminary

```

A:Molecule type: DNA
A:Residues: 1-213 <STO>
A:Cross-references: UNIPROT:Q22263; UNIPARC:UPI000017A684; GB:chr_V; PIDN:CAA94795.1; PID:
C:Genetics:
A:Gene: T0684.9
A:Map position: 5

Query Match      83.8%; Score 31; DB 2; Length 213;
Best Local Similarity 71.4%; Pred. No. 1.8e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ILAPPVP 7
   :|||||
Db 116 LLAPPAP 122

RESULT 43
F95271
hypothetical protein Sma0148 [imported] - Sinorhizobium meliloti (strain 1021) magaplasma
C:Species: Sinorhizobium meliloti
C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
C:Accession: F95271
R:Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bowe
.; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C.
Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001
A:Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium meliloti
A:Reference number: A95262; MUID:21396509; PMID:11481432
A:Accession: F95271
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-237 <KUR>
A:Cross-references: UNIPROT:Q930W6; UNIPARC:UPI00000CAF86; GB:AE006469; PIDN:AAK64736.1;
A:Experimental source: strain 1021, megaplasmaid pSymA
R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
L.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A:Reference number: A96039; MUID:21368234; PMID:11474104
A:Contents: annotation
C:Genetics:
A:Gene: Sma0148
A:Genome: plasmid

Query Match      83.8%; Score 31; DB 2; Length 237;
Best Local Similarity 57.1%; Pred. No. 1.8e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAPPVP 7
   :|||||
Db 23 VIAPPUP 29

RESULT 44
T42539
hypothetical protein - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jul-2004
C:Accession: T42539
R:Yoshioka, S.; Kato, K.; Nakai, K.; Okayama, H.; Nojima, H.
DNA Res. 4, 363-369, 1997
A:Title: Identification of open reading frames in Schizosaccharomyces pombe cDNAs.
A:Reference number: Z17323; MUID:98162722; PMID:9501991
A:Accession: T42539
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-239 <YOS>
A:Cross-references: UNIPROT:P78826; UNIPARC:UPI000006B5DC; EMBL:D89174; NID:g1749555; PI
A:Experimental source: strain PR745

Query Match      83.8%; Score 31; DB 2; Length 239;

```

```

Best Local Similarity 71.4%; Pred. No. 1.8e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ILAPPVP 7
   :|||||
Db 17 IICPPVP 23

RESULT 45
B42768
homeotic protein goosecoid - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: B42768; D40856
R:Blum, M.; Gaunt, S.J.; Cho, K.W.Y.; Steinbeisser, H.; Blumberg, B.; Bittner, D.; De Ro
Cell 69, 1097-1106, 1992
A:Title: Gastrulation in the mouse: the role of the homeobox gene goosecoid.
A:Reference number: A42768; MUID:92315328; PMID:1352187
A:Accession: B42768
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-243 <BLU>
A:Cross-references: UNIPARC:UPI00001745C0; GB:M63872
R:Blumberg, B.; Wright, C.V.E.; De Robertis, E.M.; Cho, K.W.Y.
Science 253, 194-196, 1991
A:Title: Organizer-specific homeobox genes in Xenopus laevis embryos.
A:Reference number: A40856; MUID:91305940; PMID:1677215
A:Accession: D40856
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 148-207 <BL2>
A:Cross-references: UNIPARC:UPI00001745C1; GB:M63872
C:Superfamily: homeotic protein goosecoid; homeobox homology
C:Keywords: DNA binding; homeobox; nucleus; transcription regulation
F:149-205/Domain: homeobox homology <HOX>

Query Match      83.8%; Score 31; DB 1; Length 243;
Best Local Similarity 71.4%; Pred. No. 1.8e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ILAPPVP 7
   :|||||
Db 114 VLPPVP 120

RESULT 46
I51424
homeotic protein goosecoid [similarity] - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 09-Jul-2004
C:Accession: I51424
R:Blumberg, B.; Cho, K.W.Y.; Steinbeisser, H.; DeRobertis, E.M.
Cell 67, 1111-1120, 1991
A:Title: Molecular nature of Spemann's organizer: The role of the Xenopus homeobox gene
A:Reference number: I51424; MUID:92103677; PMID:1684739
A:Accession: I51424
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-243 <BLU>
A:Cross-references: UNIPROT:P53546; UNIPARC:UPI000012BBA1; GB:M81481; NID:g214185; PIDN:
C:Superfamily: homeotic protein goosecoid; homeobox homology
C:Keywords: DNA binding; homeobox; nucleus; transcription regulation
F:149-205/Domain: homeobox homology <HOX>

Query Match      83.8%; Score 31; DB 2; Length 243;
Best Local Similarity 71.4%; Pred. No. 1.8e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ILAPPVP 7
   :|||||
Db 114 VLPPVP 120

```

RESULT 47

AP2138
 Hypothetical protein all2661 [imported] - Nostoc sp. (strain PCC 7120)
 C:Species: Nostoc sp. PCC 7120
 A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
 C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
 C:Accession: AF2138
 R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, H.; Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
 DNA Res. 8, 205-213, 2001
 A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena PCC 7120
 A:Reference number: AB1807; MUID:21595285; PMID:11759840
 A:Accession: AF2138
 A:Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-326 <KUR>
 A:Cross-references: UNIPROT:Q8YTQ2; UNIPARC:UPI00000CE4PB; GB:BA000019; PIDN:BA074360.1
 A:Experimental source: strain PCC 7120
 C:Genetics:
 A:Gene: all2661

Query Match 83.8%; Score 31; DB 2; Length 326;
 Best Local Similarity 71.4%; Pred. No. 2.5e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ILAPPVP 7
 | : | : | : |
 Db 189 ITAPPIP 195

RESULT 48

T51269
 Hypothetical protein T8M16_140 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 09-Jul-2004
 C:Accession: T51269
 R:Benes, V.; Wurbach, E.; Drzonek, H.; Ansoorge, W.; Meves, H.W.; Rudd, S.; Lemcke, K.;
 submitted to the Protein Sequence Database, August 2000
 A:Reference number: Z25346
 A:Accession: T51269
 A:Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-332 <BEN>
 A:Cross-references: UNIPROT:Q9LBS7; UNIPARC:UPI00000A48CC; EMBL:AL390921
 A:Experimental source: cultivar Columbia; BAC clone T8M16
 C:Genetics:
 A:Map position: 3
 A:Introns: 150/1
 A:Note: T8M16_140

Query Match 83.8%; Score 31; DB 2; Length 332;
 Best Local Similarity 71.4%; Pred. No. 2.6e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAPPVP 7
 | : | : | : |
 Db 83 IVAPPVP 89

RESULT 49

A96835
 Gibberellin 3 beta-hydroxylase, 29683-28215 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
 C:Accession: A96835
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, J.; Holt, T.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.-A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A66141; MUID:21016719; PMID:11130712
 A:Accession: A96835
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-347 <STO>
 A:Cross-references: UNIPROT:Q9C970; UNIPARC:UPI00000484C6; GB:AB005173; NID:96751709; PI:1
 C:Genetics:
 A:Map position: 1
 A:Gene: F516.9
 C:Superfamily: 1-aminocyclopropane-1-carboxylate oxidase

Query Match 83.8%; Score 31; DB 2; Length 347;
 Best Local Similarity 71.4%; Pred. No. 2.7e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ILAPPVP 7
 | : | : | : |
 Db 252 VTAPPVP 258

RESULT 50

T51691
 Gibberellin 3 beta-hydroxylase [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 09-Jul-2004
 C:Accession: T51691
 R:Yamaguchi, S.; Smith, M.W.; Brown, R.G.; Kamiya, Y.; Sun, T.
 Plant Cell 10, 2115-2126, 1998
 A:Title: Phytochrome regulation and differential expression of gibberellin 3beta-hydroxy
 A:Reference number: Z25425; MUID:99055200; PMID:9836749
 A:Accession: T51691
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-347 <YAM>
 A:Cross-references: UNIPROT:Q9ZT84; UNIPARC:UPI00000A0C66; EMBL:AF070937; PIDN:AAC83647.

Query Match 83.8%; Score 31; DB 2; Length 347;
 Best Local Similarity 71.4%; Pred. No. 2.7e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ILAPPVP 7
 | : | : | : |
 Db 252 VTAPPVP 258

Search completed: April 6, 2006, 09:40:04
 Job time : 24.7105 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 6, 2006, 09:23:24 ; Search time 85.1053 Seconds
(without alignments)
58.030 Million cell updates/sec

Title: US-10-632-388-296

Perfect score: 37

Sequence: 1 ILAPPVP 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Database : Uniprot 05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	37	100.0	221	2	Q5M2B3_SYPN6
2	37	100.0	370	2	Q9LGV4_ORYSA
3	37	100.0	1064	2	Q8SXB2_DROSOPHILA
4	37	100.0	1607	2	Q8IRS0_DROME
5	37	100.0	1767	2	Q24495_DROME
6	37	100.0	1767	2	Q9W4F5_DROME
7	36	97.3	161	2	Q39063_ARATH
8	36	97.3	218	2	Q6ZN89_HUMAN
9	36	97.3	461	2	Q7SBK3_NEUCR
10	35	94.6	75	2	Q6Z8U5_ORYSA
11	35	94.6	186	2	Q5D5E5_9RICK
12	35	94.6	191	2	Q4UL63_RICFE
13	35	94.6	191	2	Q73HK2_WOLPM
14	35	94.6	201	2	Q5FAP8_ANAMM
15	35	94.6	231	2	Q6W0X3_CAEBR
16	35	94.6	266	2	Q7D5X3_MYCTU
17	35	94.6	266	2	Q05853_MYCTU
18	35	94.6	266	2	Q73UQ2_MYCPA
19	35	94.6	266	2	Q7TWZ9_MYCBO
20	35	94.6	477	2	Q5WMS2_ORYSA
21	35	94.6	662	2	Q7PYU0_ANOGA
22	35	94.6	842	2	Q55ZY4_CRYNE
23	35	94.6	842	2	Q5KP92_CRYNE
24	35	94.6	990	2	Q6Y8G9_LEIME
25	35	94.6	990	2	Q4QG10_LEIMA
26	34	91.9	189	1	PAAD_RICPR
27	34	91.9	190	2	Q68WJ3_RICTY
28	34	91.9	315	2	Q62156_BURMA
29	34	91.9	341	2	Q5R192_BRARE
30	34	91.9	467	2	Q4SLJ8_TETNG
31	34	91.9	621	2	Q4IM73_GIBZE

32	34	91.9	762	2	Q8T3T1_LYTVA	08t3t1 lytechinus
33	34	91.9	768	2	Q8C585_MOUSE	08c585 mus musculus
34	34	91.9	924	2	Q73172_SNIY3	Q73172 synechocyst
35	34	91.9	959	2	Q520P7_MAGGR	Q520p7 magnaporthe
36	34	91.9	1011	2	Q52FV3_MAGGR	Q52fv3 mus musculus
37	34	91.9	1048	2	Q80TV1_MOUSE	Q80tv1 mus musculus
38	34	91.9	1291	2	Q6Y7W8_MOUSE	Q6y7w8 mus musculus
39	33	89.2	111	2	Q857Z6_9CAUD	Q857z6 mycobacteri
40	33	89.2	113	1	Y124_LISIN	Q925x3 listeria in
41	33	89.2	129	2	Q8DKV4_SYNEL	Q8dkv4 synechococ
42	33	89.2	147	2	Q9R6A1_9RHIZ	Q9r6a1 agrobacteri
43	33	89.2	164	2	Q9K3W9_STRCO	Q9k3w9 streptomyce
44	33	89.2	177	2	Q6Z9K3_ORYSA	Q6z9k3 oryza sativ
45	33	89.2	183	2	Q73TA7_MYCPA	Q73ta7 mycobacteri
46	33	89.2	205	1	HLH4_CAEEL	P34555 caenorhabdi
47	33	89.2	226	2	Q97177_RHOSH	P97177 rhodobacter
48	33	89.2	228	2	Q5FC67_CAEEL	Q5fc67 caenorhabdi
49	33	89.2	235	2	Q52WY3_BRARE	Q52wy3 brachydanio
50	33	89.2	236	2	Q89E19_BRAJA	Q89e19 bradyrhizob
51	33	89.2	237	2	Q568D3_BRARE	Q568d3 brachydanio
52	33	89.2	258	2	Q9LIH7_STRCO	Q9lih7 streptomyce
53	33	89.2	264	2	Q82C90_STRAW	Q82c90 streptomyce
54	33	89.2	281	2	Q82G67_STRAW	Q82g67 streptomyce
55	33	89.2	284	2	Q7XED3_ORYSA	Q7xed3 oryza sativ
56	33	89.2	315	2	Q63RF2_BURPS	Q63rf2 burkholderi
57	33	89.2	317	2	Q52286_9RHIZ	Q52286 agrobacteri
58	33	89.2	324	2	Q84YF5_SORBI	Q84yf5 sorghum bic
59	33	89.2	325	2	Q84YF3_SORBI	Q84yf3 sorghum bic
60	33	89.2	341	2	Q7R666_GIALA	Q7r666 giardia lam
61	33	89.2	370	2	Q8LGL0_SILDI	Q8lgl0 silene dioi
62	33	89.2	371	2	Q5K483_SILDI	Q5k483 silene dioi
63	33	89.2	375	2	Q4Q8Y4_LEIMA	Q4q8y4 leishmania
64	33	89.2	393	2	Q9EWB0_STRCO	Q9ewb0 streptomyce
65	33	89.2	418	2	Q4IBX5_GIBZE	Q4ibx5 gibberella
66	33	89.2	430	2	Q4HY43_GIBZE	Q4hy43 gibberella
67	33	89.2	444	2	Q965C9_GIALA	Q965c9 giardia lam
68	33	89.2	444	2	Q4TP09_9SPHN	Q4tp09 erythrobact
69	33	89.2	450	2	Q6MYW6_ASFFU	Q6myw6 aspergillus
70	33	89.2	452	2	Q7SCZ7_NEUCR	Q7scz7 neurospora
71	33	89.2	455	2	Q5BEB0_EMENI	Q5beb0 aspergillus
72	33	89.2	455	2	Q4WST6_ASFFU	Q4wst6 aspergillus
73	33	89.2	464	2	Q75HS1_ORYSA	Q75hs1 oryza sativ
74	33	89.2	465	2	Q54176_STRCO	Q54176 streptomyce
75	33	89.2	480	2	Q7R5N0_GIALA	Q7r5n0 giardia lam
76	33	89.2	483	2	Q5AZ77_EMENI	Q5az77 aspergillus
77	33	89.2	497	2	Q51HI0_MAGGR	Q51hi0 magnaporthe
78	33	89.2	506	2	Q6TV65_9POXV	Q6tv65 bovine papu
79	33	89.2	509	2	Q50LG1_ALTAL	Q50lg1 alternaria
80	33	89.2	516	2	Q51X61_MAGGR	Q51x61 magnaporthe
81	33	89.2	520	2	Q7F595_ORYSA	Q7f595 oryza sativ
82	33	89.2	520	2	Q9SLZ5_ORYSA	Q9slz5 oryza sativ
83	33	89.2	531	2	Q6C5V4_YARLI	Q6c5v4 yarrowia li
84	33	89.2	532	2	Q6TVJ6_9POXV	Q6tvj6 orf virus
85	33	89.2	532	2	Q6TVX6_9POXV	Q6tvx6 orf virus
86	33	89.2	535	2	Q8AYG6_RAT	Q8ayg6 rattus norv
87	33	89.2	538	2	P74506_SINY3	P74506 synechocyst
88	33	89.2	540	2	Q5JLS2_ORYSA	Q5jls2 oryza sativ
89	33	89.2	543	2	Q8YY82_ANASP	Q8yy82 anabaena sp
90	33	89.2	551	2	Q46113_DROME	Q46113 drosophila
91	33	89.2	566	2	Q5WN70_CAEBR	Q5wn70 caenorhabdi
92	33	89.2	568	2	Q81ZZ0_STRAW	Q81zz0 streptomyce
93	33	89.2	572	2	Q41JP9_GIBZE	Q41jp9 gibberella
94	33	89.2	588	2	Q5B1Q1_EMENI	Q5b1q1 aspergillus
95	33	89.2	592	2	Q7NCL0_GLOVI	Q7ncl0 gloeobacter
96	33	89.2	597	2	Q6ZH89_ORYSA	Q6zh89 oryza sativ
97	33	89.2	601	2	Q4HA07_9DEIO	Q4ha07 deinococcus
98	33	89.2	604	2	Q4S2Y2_TETNG	Q4s2y2 tetraodon n
99	33	89.2	635	2	Q67UX0_ORYSA	Q67ux0 oryza sativ
100	33	89.2	637	2	Q61VG5_EMENI	Q61vg5 emericeila
101	33	89.2	640	2	Q5ZJW0_CHICK	Q5zjw0 gallus gall
102	33	89.2	644	2	Q5NKM3_CRYNV	Q5nkm3 cryptococu
103	33	89.2	647	2	Q5R9F2_PONPY	Q5r9f2 pongo pygma
104	33	89.2	648	2	Q7PXC5_ANOGA	Q7pxc5 anopheles g

105	33	89.2	687	2	059895	PODAN	059895	podospora a	178	32	86.5	297	2	Q8LA38	ARATH	Q8LA38	arabidopsis
106	33	89.2	766	2	Q52EX7	MAGGR	Q52EX7	magnaporthe	179	32	86.5	307	2	Q587Y6	PORYE	Q587Y6	porphyra ye
107	33	89.2	768	2	Q58LD2	CAUD	Q58LD2	cyanophaga	180	32	86.5	310	2	Q527E0	MAGGR	Q527E0	magnaporthe
108	33	89.2	779	2	Q4Q991	LEIMA	Q4Q991	leishmania	181	32	86.5	317	2	Q4QBD0	LEIMA	Q4QBD0	leishmania
109	33	89.2	782	2	Q4WLT7	ASPFU	Q4WLT7	aspergillus	182	32	86.5	324	2	Q90Z55	SCOMX	Q90Z55	scophthalmu
110	33	89.2	786	2	Q4T161	TETNG	Q4T161	tetradon n	183	32	86.5	335	2	Q5LV54	SILPO	Q5LV54	silicibacte
111	33	89.2	819	1	FTSK	CAUCR	Q9A262	caulobacter	184	32	86.5	358	2	Q4KY19	ACASC	Q4KY19	acanthopagr
112	33	89.2	830	2	Q78967	NEUCR	Q78967	neurospora	185	32	86.5	374	2	Q787V1	NEUCR	Q787V1	neurospora
113	33	89.2	835	2	Q6K317	ORYZA	Q6K317	oryza sativ	186	32	86.5	402	2	Q5WQW1	CRYNE	Q5WQW1	cryptococcu
114	33	89.2	853	2	Q59GQ3	HUMAN	Q59GQ3	homo sapien	187	32	86.5	409	2	Q4H7W5	9DEIO	Q4H7W5	deinococcu
115	33	89.2	861	2	Q69JZ8	ORYSA	Q69JZ8	oryza sativ	188	32	86.5	422	2	Q6A8M0	PROAC	Q6A8M0	propionibac
116	33	89.2	870	1	MILK1	MOUSE	Q8BGT6	mus musculu	189	32	86.5	425	2	Q6ANN2	9DELIT	Q6ANN2	anaeromyxob
117	33	89.2	905	2	Q658W4	HUMAN	Q658W4	homo sapien	190	32	86.5	426	2	Q55R33	CRYNE	Q55R33	cryptococcu
118	33	89.2	920	2	Q4P159	USTILAGO	Q4P159	ustilago ma	191	32	86.5	428	2	Q5KDW3	CRYNE	Q5KDW3	cryptococcu
119	33	89.2	962	2	Q72Z33	homo sapien	Q72Z33	homo sapien	192	32	86.5	428	2	Q4RPX5	TETNG	Q4RPX5	tetradon n
120	33	89.2	977	1	MCR	TUPGB	Q29131	tupaia glis	193	32	86.5	430	2	Q8X016	NEUCR	Q8X016	neurospora
121	33	89.2	978	1	MCR	MOUSE	Q8V118	mus musculu	194	32	86.5	433	2	Q94IO9	MAIZE	Q94IO9	zea mays (m
122	33	89.2	980	1	RINJ	MOUSE	P59729	mus musculu	195	32	86.5	439	2	Q51KW7	MAGGR	Q51KW7	magnaporthe
123	33	89.2	980	2	Q80WAO	MOUSE	Q80WAO	mus musculu	196	32	86.5	460	2	Q9VM57	DROME	Q9VM57	drosophila
124	33	89.2	980	1	MCR	RAIF	P22199	rattus norv	197	32	86.5	464	2	Q52CB0	MAGGR	Q52CB0	magnaporthe
125	33	89.2	982	1	MCR	SAISC	Q9H0W8	saimiri sci	198	32	86.5	469	2	Q4RK08	TETNG	Q4RK08	tetradon n
126	33	89.2	982	2	Q4JMN28	9PRIM	Q4JMN28	saimiri bol	199	32	86.5	471	2	Q9E128	9ALPH	Q9E128	carcophithe
127	33	89.2	983	2	Q6XLI19	CALJJA	Q6XLI19	callithrix	200	32	86.5	481	2	Q9SY17	ARATH	Q9SY17	arabidopsis
128	33	89.2	984	1	MCR	HUMAN	P08235	homo sapien	201	32	86.5	497	2	Q4T3X4	TETNG	Q4T3X4	tetradon n
129	33	89.2	987	2	Q6XLI18	CALJJA	P08235	homo sapien	202	32	86.5	499	2	Q58DJ0	BOVIN	Q58DJ0	bos taurus
130	33	89.2	1092	2	Q4NN66	9DELIT	Q4NN66	anaeromyxob	203	32	86.5	510	2	Q98P15	9LAMI	Q98P15	asarina bar
131	33	89.2	1094	2	Q5B395	EMENI	Q5B395	aspergillus	204	32	86.5	518	2	Q5AZ81	EMENI	Q5AZ81	aspergillus
132	33	89.2	1214	2	Q73835	MYCPA	Q73835	mycobacteri	205	32	86.5	519	2	P93833	ARATH	P93833	arabidopsis
133	33	89.2	1252	2	Q4PC18	USTILAGO	Q4PC18	ustilago ma	206	32	86.5	524	2	Q9VUU9	DROME	Q9VUU9	drosophila
134	33	89.2	1307	2	Q9V4J6	DROME	Q9V4J6	drosophila	207	32	86.5	530	2	Q7QFM7	ANOQA	Q7QFM7	anopheles g
135	33	89.2	1329	1	GP124	MOUSE	Q91ZV8	mus musculu	208	32	86.5	557	2	Q50EX8	SPAAU	Q50EX8	sparus aura
136	33	89.2	1331	1	GP124	HUMAN	Q9GPE1	homo sapien	209	32	86.5	563	2	Q4PC28	USTMA	Q4PC28	ustilago ma
137	33	89.2	1337	2	Q4Q3N2	LEIMA	Q4Q3N2	leishmania	210	32	86.5	567	2	Q6CAG9	YARLI	Q6CAG9	yarrowia li
138	33	89.2	1367	2	Q20120	CABEL	Q20120	caenorhabdi	211	32	86.5	575	2	Q5NVA9	PONPY	Q5NVA9	pongo pygma
139	33	89.2	1413	2	Q582N5	9TRYLP	Q582N5	trypanosoma	212	32	86.5	583	2	Q50H40	SPAAU	Q50H40	sparus aura
140	33	89.2	1489	2	Q4PE83	USTILAGO	Q4PE83	ustilago ma	213	32	86.5	590	2	Q68U44	SCOMX	Q68U44	scophthalmu
141	33	89.2	1725	2	Q621F2	BURMA	Q621F2	burkholderi	214	32	86.5	597	2	Q9C586	ARATH	Q9C586	arabidopsis
142	33	89.2	1814	2	Q6WAY7	PEA	Q6WAY7	pisum sativ	215	32	86.5	602	2	Q7T189	SILME	Q7T189	silurus mer
143	33	89.2	1841	2	Q63VY6	BURPS	Q63VY6	burkholderi	216	32	86.5	605	2	Q4R7G9	MACFA	Q4R7G9	macaca fasc
144	33	89.2	2137	2	Q84BQ6	9PSED	Q84BQ6	pseudomonas	217	32	86.5	609	1	PECK	STRO	PECK	strepomyce
145	33	89.2	2147	2	Q91950	9PSEPU	Q91950	pseudomonas	218	32	86.5	611	2	Q8W420	ARATH	Q8W420	arabidopsis
146	33	89.2	2348	2	Q521P8	NOCPA	Q521P8	nocardia fa	219	32	86.5	613	2	Q9H9G8	HUMAN	Q9H9G8	homo sapien
147	33	89.2	2518	2	Q8GBX7	POLCB	Q8GBX7	polyangium	220	32	86.5	623	2	Q6UAQ0	TETNG	Q6UAQ0	tetradon n
148	33	89.2	6310	2	Q88PP2	PSEPK	Q88PP2	pseudomonas	221	32	86.5	633	2	Q90Z56	SCOMX	Q90Z56	scophthalmu
149	33	89.2	7599	2	Q41K83	GIBZE	Q41K83	gibberella	222	32	86.5	640	2	Q8AXA5	ACASC	Q8AXA5	acanthopagr
150	33	89.2	7962	2	Q10465	HUMAN	Q10465	homo sapien	223	32	86.5	642	2	Q23673	ARATH	Q23673	arabidopsis
151	32	86.5	51	2	Q865J2	LEMCA	Q865J2	lemur catta	224	32	86.5	647	2	Q71MN2	SPAAU	Q71MN2	sparus aura
152	32	86.5	57	2	Q865J3	LEMVA	Q865J3	lemur varie	225	32	86.5	657	2	Q60H17	ONCMA	Q60H17	oncorhynch
153	32	86.5	57	2	Q6H3Y1	ORYSA	Q6H3Y1	oryza sativ	226	32	86.5	667	2	Q4X0B7	ASPFU	Q4X0B7	aspergillus
154	32	86.5	68	2	Q51ICO	MAGGR	Q51CO	magnaporthe	227	32	86.5	693	2	Q8WVW4	HUMAN	Q8WVW4	homo sapien
155	32	86.5	101	2	Q8DZ9	STRA5	Q8DZ9	streptococc	228	32	86.5	708	2	Q8V616	9VIRU	Q8V616	trichomonas
156	32	86.5	101	2	Q8E5G8	STRA3	Q8E5G8	streptococc	229	32	86.5	720	2	Q9CA22	ARATH	Q9CA22	arabidopsis
157	32	86.5	109	2	Q7WQ05	BORBR	Q7WQ05	bordebella	230	32	86.5	732	2	Q53T46	HUMAN	Q53T46	homo sapien
158	32	86.5	109	2	Q98MR9	RHTLO	Q98MR9	rhizobium l	231	32	86.5	777	2	Q84VX4	ARATH	Q84VX4	arabidopsis
159	32	86.5	115	2	Q61LY5	DROME	Q61LY5	drosophila	232	32	86.5	836	2	Q4SPR0	TETNG	Q4SPR0	tetradon n
160	32	86.5	121	2	Q958K3	RANAU	Q958K3	rana aurora	233	32	86.5	864	2	Q4FYD4	LEIMA	Q4FYD4	leishmania
161	32	86.5	123	2	Q69U83	ORYSA	Q69U83	oryza sativ	234	32	86.5	872	2	Q5SRQ8	CRYNE	Q5SRQ8	cryptococcu
162	32	86.5	125	2	Q5N221	SYNPP6	Q5N221	synchococc	235	32	86.5	879	2	Q5KG63	CRYNE	Q5KG63	cryptococcu
163	32	86.5	127	2	Q60H16	ONCKI	Q60H16	oncorhynch	236	32	86.5	885	2	Q6P553	MOUSE	Q6P553	mus musculu
164	32	86.5	128	2	Q5NAP5	ORYSA	Q5NAP5	oryza sativ	237	32	86.5	885	2	Q4H2K1	CIOIN	Q4H2K1	ciona intes
165	32	86.5	157	2	Q8N1G8	HUMAN	Q8N1G8	homo sapien	238	32	86.5	894	2	Q5YT66	NOCPA	Q5YT66	nocardia fa
166	32	86.5	164	2	Q60H15	ONCTS	Q60H15	oncorhynch	239	32	86.5	911	1	PTSK	SHEON	PTSK	shewanella
167	32	86.5	165	2	Q61UG9	DROME	Q61UG9	drosophila	240	32	86.5	1048	2	Q9H0E3	HUMAN	Q9H0E3	homo sapien
168	32	86.5	187	2	Q5GTLO	WOLTRA	Q5GTLO	wolbachia s	241	32	86.5	1056	2	Q8B1H0	MOUSE	Q8B1H0	mus musculu
169	32	86.5	191	2	Q7P8V3	RICSI	Q7P8V3	rickettsia	242	32	86.5	1098	2	Q6NZP5	MOUSE	Q6NZP5	mus musculu
170	32	86.5	191	2	Q92HH0	RICCN	Q92HH0	rickettsia	243	32	86.5	1100	2	Q9NV82	HUMAN	Q9NV82	homo sapien
171	32	86.5	194	2	Q89D15	BRATA	Q89D15	bradyrhizob	244	32	86.5	1103	2	Q59FN7	HUMAN	Q59FN7	homo sapien
172	32	86.5	200	2	Q86YU6	HUMAN	Q86YU6	homo sapien	245	32	86.5	1148	2	Q789L4	NEUCR	Q789L4	neurospora
173	32	86.5	206	2	Q89PF2	BRATA	Q89PF2	bradyrhizob	246	32	86.5	1202	2	Q5B6T7	EMENI	Q5B6T7	aspergillus
174	32	86.5	210	2	Q9YCV5	AERPE	Q9YCV5	aeropyrum p	247	32	86.5	1298	2	Q7Z312	HUMAN	Q7Z312	homo sapien
175	32	86.5	226	2	Q6L5A6	ORYSA	Q6L5A6	oryza sativ	248	32	86.5	1299	2	Q6Y7W6	HUMAN	Q6Y7W6	homo sapien
176	32	86.5	228	1	Y132	NPVOP	P24080	orgyia pseu	249	32	86.5	1299	2	Q7Z228	ARATH	Q7Z228	homo sapien
177	32	86.5	233	2	Q5YUT5	NOCPA	Q5YUT5	nocardia fa	250	32	86.5	1309	2	Q8GZA2	ARATH	Q8GZA2	arabidopsis

251	32	86.5	1309	2	Q9LUG9_ARATH	Q9LUG9_arabidopsis	324	31	83.8	237	2	Q930W6_RHIME	Q930w6_rhizobium m
252	32	86.5	1316	2	Q41PP2_GIBBEZ	Q41pp2_gibberella	325	31	83.8	239	1	YABF_SCHPO	Q09814_schizosacch
253	32	86.5	1320	2	Q5A5B8_CANAL	Q5a5b8_candida alb	326	31	83.8	239	1	P78826_SCHPO	P78826_schizosacch
254	32	86.5	1329	2	Q75137_HUMAN	Q75137_homo sapien	327	31	83.8	243	1	GSCA_XENLA	P29454_xenopus lae
255	32	86.5	1417	2	Q6C2D2_YARLI	Q6c2d2_yarrowia li	328	31	83.8	243	1	GSCB_XENLA	P53546_xenopus lae
256	32	86.5	1448	2	Q4REE9_TERNG	Q4ree9_tetradodon n	329	31	83.8	243	2	Q68F73_XENLA	Q68f73_xenopus lae
257	32	86.5	1610	2	Q22173_CAEBL	Q22173_caenorhabdi	330	31	83.8	243	2	Q4LDP9_XENTR	Q4ldp9_xenopus tro
258	32	86.5	1612	2	Q6B7M5_DEBHA	Q6b7m5_debaryomyce	331	31	83.8	251	2	Q7NNL3_GLOVI	Q7nnl3_gloeobacter
259	32	86.5	1626	2	Q9RY17_DEIRA	Q9ry17_deinococcus	332	31	83.8	265	2	Q9X551_CORGL	Q9x551_corynebacte
260	32	86.5	1717	2	Q21875_CAEBL	Q21875_caenorhabdi	333	31	83.8	270	2	Q75SV6_ONCMA	Q75sv6_oncorynchu
261	32	86.5	1960	2	Q7R4L8_GIALA	Q7r4l8_giardia lam	334	31	83.8	274	2	Q73F23_WOLPM	Q73f23_wolbachia p
262	32	86.5	2066	2	Q923T8_PSESX	Q923t8_pseudomonas	335	31	83.8	275	2	Q5Z4H2_ORYSA	Q5z4h2_oryza sativ
263	32	86.5	2066	2	Q87M69_PSESM	Q87m69_pseudomonas	336	31	83.8	277	2	Q8RST7_EWCH	Q8rst7_erwinia chr
264	32	86.5	2106	2	Q6C0Y6_YARLI	Q6c0y6_yarrowia li	337	31	83.8	291	2	Q51TQ6_MAGGR	Q51tq6_magnaporthe
265	32	86.5	2262	2	Q6WAY3_FEA	Q6way3_pisum sativ	338	31	83.8	294	2	Q6J243_9PLEO	Q6j243_phaeospaer
266	32	86.5	3006	2	Q9Y4D8_HUMAN	Q9y4d8_homo sapien	339	31	83.8	295	2	Q8J244_9PLEO	Q8j244_phaeospaer
267	32	86.5	3530	1	MYO15_HUMAN	Q9ukn7_homo sapien	340	31	83.8	298	2	Q4ITE0_AZOV1	Q4ite0_azotobacter
268	32	86.5	34350	2	Q8WZ42_HUMAN	Q8wz42_homo sapien	341	31	83.8	303	2	Q4K9B9_PSEF5	Q4k9b9_pseudomonas
269	31	83.8	52	2	Q6K2I5_ORYSA	Q6k2i5_oryza sativ	342	31	83.8	311	2	Q41YM5_AZOV1	Q41ym5_azotobacter
270	31	83.8	110	2	Q5V4P1_HALMA	Q5v4p1_haloarcula	343	31	83.8	321	2	Q4SP11_TSTNG	Q4sp11_tetradodon n
271	31	83.8	123	2	Q4LXX5_9BURK	Q4lxx5_burkholderi	344	31	83.8	322	2	Q6Q1A6_9DELA	Q6q1a6_human t-lym
272	31	83.8	127	2	Q53WJ6_ORYSA	Q53wj6_oryza sativ	345	31	83.8	326	2	Q8YTQ2_ANASP	Q8ytq2_anabaena sp
273	31	83.8	133	2	Q59PD4_CANAL	Q59pd4_candida alb	346	31	83.8	328	2	Q4NJT6_9MICC	Q4njt6_burkholderi
274	31	83.8	133	2	Q96VL8_CANAL	Q96vl8_candida alb	347	31	83.8	328	2	Q4M1A7_9BURK	Q4mla7_burkholderi
275	31	83.8	141	2	Q6QUJ6_9RHIZ	Q6quj6_hyphomicrob	348	31	83.8	329	2	Q9M5K1_EUPES	Q9m5k1_euphorbia e
276	31	83.8	142	2	Q6QUJ4_9RHIZ	Q6quj4_hyphomicrob	349	31	83.8	332	2	Q9LES7_ARATH	Q9les7_arabidopsis
277	31	83.8	142	2	Q6QUJ5_9RHIZ	Q6quj5_hyphomicrob	350	31	83.8	333	2	Q8KGB5_CHLTE	Q8kgb5_chlorobium
278	31	83.8	162	2	Q9A3B4_CAUCR	Q9a3b4_caulobacter	351	31	83.8	335	2	Q8B2W8_9PARA	Q8b2w8_measles vir
279	31	83.8	164	2	Q70R68_9HIV1	Q70r68_human immun	352	31	83.8	337	2	Q5EMQ9_9HIV1	Q5emq9_human immun
280	31	83.8	171	2	Q6EEU0_9HIV1	Q6eeu0_human immun	353	31	83.8	347	2	Q9C970_ARATH	Q9c970_arabidopsis
281	31	83.8	171	2	Q6EEU1_9HIV1	Q6eeu1_human immun	354	31	83.8	347	2	Q9ZT84_ARATH	Q9zt84_arabidopsis
282	31	83.8	178	1	DEF2_PSEPK	Q8eeat_pseudomonas	355	31	83.8	352	2	Q8FQQ9_9HIV1	Q8fqq9_allium cepa
283	31	83.8	179	1	DEF2_PSESM	Q886l1_pseudomonas	356	31	83.8	352	2	Q84TMO_9HIV1	Q84tmo_allium cepa
284	31	83.8	179	2	Q4ZPW1_PSESY	Q4zpw1_pseudomonas	357	31	83.8	361	2	Q57Z20_9TRYR	Q57z20_trypanosoma
285	31	83.8	179	2	Q4K7V5_PSEF5	Q4k7v5_pseudomonas	358	31	83.8	365	2	Q9WDB5_9HIV1	Q9wdb5_human immun
286	31	83.8	179	2	Q9OCM9_9HIV1	Q9ocm9_human immun	359	31	83.8	366	2	Q61852_CAEBL	Q61852_caenorhabdi
287	31	83.8	185	2	Q6UDP6_9HIV1	Q6udp6_human immun	360	31	83.8	374	2	Q7ZP47_9HIV1	Q7zp47_human immun
288	31	83.8	189	2	Q5WST0_LEGPL	Q5wst0_legionella	361	31	83.8	379	2	Q4QHB1_LEIMA	Q4qhb1_leishmania
289	31	83.8	189	2	Q5X104_LEGPA	Q5x104_legionella	362	31	83.8	379	2	Q6N3A2_RHOPA	Q6n3a2_rhodospseudo
290	31	83.8	189	2	Q5ZRJ8_LEGPH	Q5zrj8_legionella	363	31	83.8	391	2	Q61V50_CABBR	Q61v50_caenorhabdi
291	31	83.8	193	2	Q70RC0_9HIV1	Q70rc0_human immun	364	31	83.8	392	2	Q4LQ85_9BURK	Q4lq85_burkholderi
292	31	83.8	193	2	Q9DVF3_9HIV1	Q9dvf3_human immun	365	31	83.8	396	2	Q87677_9PSEU	Q87677_amycolatops
293	31	83.8	194	2	Q7SKT2_9HIV1	Q7skt2_human immun	366	31	83.8	402	2	Q5S972_9SYN3	Q5s972_synecocyst
294	31	83.8	195	2	Q7ZNJ2_9HIV1	Q7znj2_human immun	367	31	83.8	407	2	Q7NUV1_CHRVO	Q7nuv1_chromobacte
295	31	83.8	195	2	Q9DVA5_9HIV1	Q9dva5_human immun	368	31	83.8	407	2	Q4ZIU5_9HEPC	Q4zius_hepatitis c
296	31	83.8	199	2	Q22262_CAEBL	Q22262_caenorhabdi	369	31	83.8	410	1	ATIN_VZVD	P09265_varicella-z
297	31	83.8	199	2	Q7SKS5_9HIV1	Q7ske5_human immun	370	31	83.8	410	2	Q6QCP5_HHV3	Q6qcp5_human herpe
298	31	83.8	200	2	Q9DV90_9HIV1	Q9dv90_human immun	371	31	83.8	410	2	Q4TOW5_HHV3	Q4tqw5_human herpe
299	31	83.8	204	2	Q7ZNX3_9HIV1	Q7znx3_human immun	372	31	83.8	410	2	Q4JQY8_HHV3	Q4jqy8_human herpe
300	31	83.8	206	2	Q70R96_9HIV1	Q70r96_human immun	373	31	83.8	411	2	Q5VJS2_XENLA	Q5vj82_xenopus lae
301	31	83.8	208	2	Q8IP43_DROME	Q8ip43_drosophila	374	31	83.8	420	2	Q8DH06_XYNEL	Q8dh06_synecococc
302	31	83.8	210	2	Q9X2X7_RHIME	Q9x2x7_rhizobium m	375	31	83.8	421	2	Q6AZN0_XENLA	Q6azn0_xenopus lae
303	31	83.8	212	2	Q22263_CAEBL	Q22263_caenorhabdi	376	31	83.8	422	2	Q4ID18_GIBBEZ	Q4id18_gibberella
304	31	83.8	213	2	Q8AD01_9HIV1	Q8ad01_human immun	377	31	83.8	423	2	Q5QIE2_CHICK	Q5qie2_gallus gall
305	31	83.8	214	2	Q5SMD6_THET8	Q5smd6_thermus the	378	31	83.8	425	2	Q6NZZ9_BRARE	Q6nzz9_brachydanio
306	31	83.8	214	2	Q7Z8H0_THET2	Q7z8h0_thermus the	379	31	83.8	427	2	Q4LI48_9BURK	Q4li48_burkholderi
307	31	83.8	214	2	Q8XPW9_RALSO	Q8xpw9_raletonia s	380	31	83.8	429	2	Q90CU2_9HIV1	Q90cu2_human immun
308	31	83.8	214	2	Q70QK9_9HIV1	Q70qk9_human immun	381	31	83.8	439	2	Q75HN4_ORYSA	Q75hn4_oryza sativ
309	31	83.8	217	2	Q4UZP3_XANCP	Q4uzf3_xanthomonas	382	31	83.8	445	2	Q8N597_HUMAN	Q8n597_homo sapien
310	31	83.8	217	2	Q8FD67_XANCP	Q8fd67_xanthomonas	383	31	83.8	445	2	Q9H8K7_HUMAN	Q9h8k7_homo sapien
311	31	83.8	217	2	Q9YXK4_9HIV1	Q9yxx4_human immun	384	31	83.8	457	2	Q4QUC9_9HIV1	Q4quc9_human immun
312	31	83.8	218	2	Q7ZBC7_9HIV1	Q7zbc7_human immun	385	31	83.8	459	2	Q4QTX1_9HIV1	Q4qtx1_human immun
313	31	83.8	219	2	Q7ZP23_9HIV1	Q7zp23_human immun	386	31	83.8	459	2	Q4QTW9_9HIV1	Q4qtw9_human immun
314	31	83.8	220	2	Q90R18_9HIV1	Q90r18_human immun	387	31	83.8	460	2	Q4QTY5_9HIV1	Q4qty5_human immun
315	31	83.8	220	2	Q9WLR8_9HIV1	Q9wlr8_human immun	388	31	83.8	462	2	Q5B0B8_EMENI	Q5b0b8_aspergillus
316	31	83.8	220	2	Q9YXJ3_9HIV1	Q9yxj3_human immun	389	31	83.8	462	2	Q4QUF4_9HIV1	Q4quf4_human immun
317	31	83.8	220	2	Q9YXK5_9HIV1	Q9yxk5_human immun	390	31	83.8	462	2	Q4QUD1_9HIV1	Q4qud1_human immun
318	31	83.8	220	2	Q7ZMB5_9HIV1	Q7zmb5_human immun	391	31	83.8	462	2	Q4QUA3_9HIV1	Q4qua3_human immun
319	31	83.8	221	2	Q9QL77_9HIV1	Q9ql77_human immun	392	31	83.8	462	2	Q4QU99_9HIV1	Q4qu99_human immun
320	31	83.8	224	2	Q7ZMD7_9HIV1	Q7zmd7_human immun	393	31	83.8	462	2	Q4QU83_9HIV1	Q4qu83_human immun
321	31	83.8	224	2	Q90R25_9HIV1	Q90r25_human immun	394	31	83.8	462	2	Q4QU73_9HIV1	Q4qu73_human immun
322	31	83.8	225	2	Q9YXJ9_9HIV1	Q9yxj9_human immun	395	31	83.8	462	2	Q4QU22_9HIV1	Q4qu22_human immun
323	31	83.8	228	2	Q90QY2_9HIV1	Q90qy2_human immun	396	31	83.8	462	2	Q4QU20_9HIV1	Q4qu20_human immun

397	31	83.8	462	2	Q4QU18_9HIV1	Q4qu18 human immun	470	31	83.8	594	2	Q5MKL1_ONCMV	Q5mk11 oncorhynchu
398	31	83.8	462	2	Q4QU12_9HIV1	Q4qu12 human immun	471	31	83.8	595	2	Q50W21_ONCMA	Q50w21 oncorhynchu
399	31	83.8	462	2	Q4QU08_9HIV1	Q4qu08 human immun	472	31	83.8	595	2	Q90YN9_ONCKI	Q90yn9 oncorhynchu
400	31	83.8	462	2	Q4QU06_9HIV1	Q4qu06 human immun	473	31	83.8	608	2	Q8FQH7_COREF	Q8fqh7 corynebacte
401	31	83.8	462	2	Q4QT28_9HIV1	Q4qt28 human immun	474	31	83.8	619	2	Q23363_CABEL	Q23363 caenorhabdi
402	31	83.8	462	2	Q4QT24_9HIV1	Q4qt24 human immun	475	31	83.8	620	2	Q7XVX3_ORYSA	Q7xvx3 oryza sativ
403	31	83.8	462	2	Q4QT22_9HIV1	Q4qt22 human immun	476	31	83.8	629	2	Q8V089_RALSO	Q8v089 ralstonia s
404	31	83.8	462	2	Q4QT20_9HIV1	Q4qt20 human immun	477	31	83.8	629	2	Q6L631_ANGJA	Q6l631 anguilla ja
405	31	83.8	462	2	Q4QTX9_9HIV1	Q4qtx9 human immun	478	31	83.8	644	2	Q6AF03_LEIXX	Q6af03 leifsonia x
406	31	83.8	463	2	Q4QUA1_9HIV1	Q4qua1 human immun	479	31	83.8	658	2	Q4WXC8_ASFPF	Q4wxc8 aspergillus
407	31	83.8	463	2	Q4QU97_9HIV1	Q4qu97 human immun	480	31	83.8	660	2	Q9V525_DROME	Q9v525 drosophila
408	31	83.8	463	2	Q4QU95_9HIV1	Q4qu95 human immun	481	31	83.8	674	2	Q8U586_AGR75	Q8u586 agrobacteri
409	31	83.8	463	2	Q4QU69_9HIV1	Q4qu69 human immun	482	31	83.8	680	2	Q729D8_DESVH	Q729d8 desulfovibr
410	31	83.8	463	2	Q4QU53_9HIV1	Q4qu53 human immun	483	31	83.8	697	2	Q5AWZ1_ENENI	Q5awz1 aspergillus
411	31	83.8	463	2	Q4QU47_9HIV1	Q4qu47 human immun	484	31	83.8	703	2	Q52CN4_MAGGR	Q52cn4 magnaporthe
412	31	83.8	463	2	Q4QU38_9HIV1	Q4qu38 human immun	485	31	83.8	720	1	V919_BPAPS	Q91b9 bacterioph
413	31	83.8	463	2	Q4QW5_9HIV1	Q4qw5 human immun	486	31	83.8	722	2	Q7S563_NEUCR	Q7s563 neurospora
414	31	83.8	469	2	Q58PX1_9HIV1	Q58px1 human immun	487	31	83.8	724	2	Q8X0X3_NEUCR	Q8x0x3 neurospora
415	31	83.8	470	2	Q9UVH4_SACPS	Q9uvh4 saccharomyc	488	31	83.8	728	2	Q8DJG9_SYNEL	Q8djg9 synechococc
416	31	83.8	474	2	Q4QU93_9HIV1	Q4qu93 human immun	489	31	83.8	729	2	Q8S6N4_ORYSA	Q8sen4 oryza sativ
417	31	83.8	474	2	Q4QU87_9HIV1	Q4qu87 human immun	490	31	83.8	738	2	Q4PBG2_USTMA	Q4pbg2 ustilago ma
418	31	83.8	474	2	Q4QU75_9HIV1	Q4qu75 human immun	491	31	83.8	748	2	Q5SMI2_CRYNE	Q5smi2 cryptococcu
419	31	83.8	474	2	Q4QU65_9HIV1	Q4qu65 human immun	492	31	83.8	760	2	Q5NB60_ORYSA	Q5nb60 oryza sativ
420	31	83.8	474	2	Q4QU63_9HIV1	Q4qu63 human immun	493	31	83.8	826	2	Q96LM4_DROME	Q96lm4 drosophila
421	31	83.8	474	2	Q4QU51_9HIV1	Q4qu51 human immun	494	31	83.8	847	2	Q69GT0_9HIV1	Q69gt0 human immun
422	31	83.8	474	2	Q4QU49_9HIV1	Q4qu49 human immun	495	31	83.8	851	2	Q7IAX6_9HIV1	Q7lax6 human immun
423	31	83.8	474	2	Q4QU43_9HIV1	Q4qu43 human immun	496	31	83.8	854	2	Q9WSB4_9HIV1	Q9wsb4 human immun
424	31	83.8	474	2	Q4QU41_9HIV1	Q4qu41 human immun	497	31	83.8	854	2	Q9YV14_9HIV1	Q9yv14 human immun
425	31	83.8	474	2	Q4QU26_9HIV1	Q4qu26 human immun	498	31	83.8	856	2	Q8JBY6_9HIV1	Q8jby6 human immun
426	31	83.8	474	2	Q4QU24_9HIV1	Q4qu24 human immun	499	31	83.8	856	2	Q9WLH6_9HIV1	Q9wlh6 human immun
427	31	83.8	476	2	Q4QTM1_9HIV1	Q4qtm1 human immun	500	31	83.8	858	2	Q4QWS7_9HIV1	Q4qws7 human immun
428	31	83.8	476	2	Q4QTV9_9HIV1	Q4qtv9 human immun	501	31	83.8	859	2	Q8JBW8_9HIV1	Q8jbw8 human immun
429	31	83.8	476	2	Q4QTV7_9HIV1	Q4qtv7 human immun	502	31	83.8	860	2	Q93091_9HIV1	Q93091 human immun
430	31	83.8	476	2	Q4QTV5_9HIV1	Q4qtv5 human immun	503	31	83.8	863	2	Q6UPP6_9HIV1	Q6upf6 human immun
431	31	83.8	476	2	Q4QTV9_9HIV1	Q4qtv9 human immun	504	31	83.8	864	2	Q5MYV2_SIVCV	Q5myv2 chimpanzee
432	31	83.8	477	2	Q4QTV1_9HIV1	Q4qtv1 human immun	505	31	83.8	864	2	Q6UEP8_9HIV1	Q6uep8 human immun
433	31	83.8	481	2	Q4QU71_9HIV1	Q4qu71 human immun	506	31	83.8	864	2	Q998H1_9HIV1	Q998h1 human immun
434	31	83.8	481	2	Q4QTM7_9HIV1	Q4qtm7 human immun	507	31	83.8	865	2	Q9IV12_9HIV1	Q9iv12 human immun
435	31	83.8	481	2	Q4QTV3_9HIV1	Q4qtv3 human immun	508	31	83.8	866	2	Q90D24_9HIV1	Q90d24 human immun
436	31	83.8	481	2	Q4QTV3_9HIV1	Q4qtv3 human immun	509	31	83.8	876	2	Q79670_9HIV1	Q79670 human immun
437	31	83.8	481	2	Q4QTU1_9HIV1	Q4qtu1 human immun	510	31	83.8	877	1	WEEL_SCHPO	P07527 schizosacch
438	31	83.8	482	2	Q52GL7_MAGGR	Q52gl7 magnaporthe	511	31	83.8	878	2	Q8JBD5_9HIV1	Q8jbd5 human immun
439	31	83.8	486	1	EGP4_HUMAN	Q05215 homo sapien	512	31	83.8	880	2	Q8PPS5_XANAC	Q8ppe5 xanthomonas
440	31	83.8	489	2	Q5CF82_CRYHO	Q5cf82 cryptospori	513	31	83.8	900	2	Q8X0N9_NEUCR	Q8xon9 neurospora
441	31	83.8	492	2	Q72V97_BRARE	Q72v97 brachydanio	514	31	83.8	919	2	Q51Q81_MAGGR	Q51q81 magnaporthe
442	31	83.8	492	2	Q98T96_BRARE	Q98t96 brachydanio	515	31	83.8	927	2	Q881W9_PBSBM	Q881w9 pseudomonas
443	31	83.8	492	2	Q85091_9DELA	Q85091 simian t-ly	516	31	83.8	937	2	Q5GVJ8_XANOR	Q5gvj8 xanthomonas
444	31	83.8	496	2	Q5SHC3_THET8	Q5shc3 thermus the	517	31	83.8	958	2	Q77274_DROME	Q77274 drosophila
445	31	83.8	496	2	Q72HN1_THET2	Q72hn1 thermus the	518	31	83.8	968	2	Q6ZIW1_ORYSA	Q6ziw1 oryza sativ
446	31	83.8	498	2	Q419A0_GIBZE	Q419a0 gibberella	519	31	83.8	978	2	Q9W4J0_DROME	Q9w4j0 drosophila
447	31	83.8	508	2	Q90CU3_9HIV1	Q90cu3 human immun	520	31	83.8	993	2	Q73US5_MYCPA	Q73us5 mycobacteri
448	31	83.8	510	2	Q55XN9_CRYNE	Q55xn9 cryptococcu	521	31	83.8	1016	2	Q6ZTQ8_HUMAN	Q6ztq8 homo sapien
449	31	83.8	510	2	Q5KM97_CRYNE	Q5km97 cryptococcu	522	31	83.8	1016	2	Q9W4I9_DROME	Q9w4i9 drosophila
450	31	83.8	517	2	Q7VSA5_BORPE	Q7vsa5 bordetella	523	31	83.8	1037	2	Q5CWR7_CRYPV	Q5cwr7 cryptospori
451	31	83.8	522	2	Q5UL31_ORYSA	Q5jl31 oryza sativ	524	31	83.8	1037	2	Q7YVY7_CRYPV	Q7yvyn7 cryptospori
452	31	83.8	524	2	Q9AJP0_STRGR	Q9ajp0 streptomyce	525	31	83.8	1038	2	Q875P1_SACKL	Q875p1 saccharomyc
453	31	83.8	524	2	Q7WC73_BORPA	Q7wc73 bordetella	526	31	83.8	1052	1	BUB1B_MOUSE	Q92180 mus musculu
454	31	83.8	524	2	Q7WQ77_BORBR	Q7wq77 bordetella	527	31	83.8	1052	2	Q8K2B5_MOUSE	Q8k2e5 mus musculu
455	31	83.8	535	2	Q91EF2_9HIV1	Q91ef2 human immun	528	31	83.8	1085	1	BUB1_HUMAN	Q43683 homo sapien
456	31	83.8	543	2	Q6FP63_CANGA	Q6fp63 candida gla	529	31	83.8	1085	2	Q53QEA_HUMAN	Q53qe4 homo sapien
457	31	83.8	553	2	Q9W4L5_DROME	Q9w4l5 drosophila	530	31	83.8	1117	2	Q5KN38_CRYNE	Q5kn38 cryptococcu
458	31	83.8	557	2	Q5ZHN9_CHICK	Q5zhn9 gallus gall	531	31	83.8	1117	2	Q5SYR7_CRYNE	Q5syrr7 cryptococcu
459	31	83.8	558	2	Q7SHL5_NEUCR	Q7shl5 neurospora	532	31	83.8	1137	2	Q4N143_THEPA	Q4n143 thellaria p
460	31	83.8	565	2	Q8YB80_BRUME	Q8yb80 bruceella me	533	31	83.8	1189	2	Q4S1J4_TETNG	Q4s1j4 tetraodon n
461	31	83.8	565	2	Q8CE13_MOUSE	Q8ce13 mus musculu	534	31	83.8	1201	2	Q51PJ1_MAGGR	Q51pj1 magnaporthe
462	31	83.8	567	2	Q95RZ8_DROME	Q95rz8 drosophila	535	31	83.8	1228	2	Q5AGD7_CANAL	Q5agd7 candida alb
463	31	83.8	573	2	Q8SQM0_BRAJA	Q8sgm0 bradyrhizob	536	31	83.8	1230	1	STE20_CANAL	Q92212 candida alb
464	31	83.8	576	2	Q91EF1_9HIV1	Q91ef1 human immun	537	31	83.8	1230	1	Q13431_CANAL	Q13431 candida alb
465	31	83.8	579	2	Q4S880_TETNG	Q4s880 tetraodon n	538	31	83.8	1239	1	M4K4_HUMAN	Q45819 homo sapien
466	31	83.8	591	2	Q5OH41_ONCMY	Q5oh41 oncorhynchu	539	31	83.8	1249	2	Q4P3M6_USTMA	Q4p3m6 ustilago ma
467	31	83.8	593	2	Q512A8_ONCMY	Q512a8 oncorhynchu	540	31	83.8	1260	2	Q53TX8_HUMAN	Q53tx8 homo sapien
468	31	83.8	593	2	Q90YP0_ONCKI	Q90yp0 oncorhynchu	541	31	83.8	1262	2	Q9N3T9_CABEL	Q9n3t9 caenorhabdi
469	31	83.8	594	2	Q5J7P6_SALSA	Q5j7p6 salmo salar	542	31	83.8	1302	2	Q9W4I8_DROME	Q9w4i8 drosophila

543	31	83.8	1306	2	077273_DROME	077273_drosophila	616	30	81.1	203	2	Q57E20_BRUAB	Q57E20 bruceella ab
544	31	83.8	1322	2	Q5B131_DROME	Q5b131_drosophila	617	30	81.1	203	2	Q8GLJ2_BRUSU	Q8glj2 bruceella au
545	31	83.8	1325	2	Q871Q0_NEUCR	Q871q0_neurospora	618	30	81.1	203	2	Q8YGD0_BRUME	Q8ygd0 bruceella me
546	31	83.8	1506	2	Q7PQLO_TRYBB	Q7pp10_anopheles g	619	30	81.1	203	2	Q4JLUI_GQBNI	Q4jlu1 goblins nige
547	31	83.8	1530	1	RPCL_TRYBB	P08968_trypanosoma	620	30	81.1	206	2	Q5H465_XANOR	Q5h465 xanthomonas
548	31	83.8	1547	1	RRPO_PNW	P20951_papaya mosa	621	30	81.1	207	2	Q5NBU6_MOUSE	Q5nbu6 mus musculu
549	31	83.8	1567	2	Q6CB83_YARLI	Q6cb83_yarrowia li	622	30	81.1	209	2	Q9B847_RHILO	Q9b847 rhizobium l
550	31	83.8	1601	2	Q54190_DICDI	Q54190_dictyosteli	623	30	81.1	212	2	Q9A7C0_CAUCR	Q9a7c0 caulobacter
551	31	83.8	1611	2	Q42854_SCHPO	Q42854_schizosacch	624	30	81.1	217	2	Q7X065_9BACT	Q7x065 uncultured
552	31	83.8	1673	2	Q9VQI9_DROME	Q9vqi9_drosophila	625	30	81.1	218	2	Q9MBH9_9VIRU	Q9mbh9 bacterioph
553	31	83.8	1687	2	Q4P9Z7_USTMA	Q4p9z7_ustilago ma	626	30	81.1	218	2	Q8LMV5_ORISA	Q8lmv5 oryza sativ
554	31	83.8	1701	2	Q521T0_NOCFA	Q521t0_nocardia fa	627	30	81.1	220	2	Q88461_9HERP	Q88461 stealth vir
555	31	83.8	1861	2	Q4WRV0_ASPFU	Q4wrv0_aspergillus	628	30	81.1	221	2	Q50523_METTF	Q50523 methanobact
556	31	83.8	2368	2	Q4S4K5_TETNG	Q4s4k5_tetradoon n	629	30	81.1	227	2	Q6K4U0_ORISA	Q6k4u0 oryza sativ
557	31	83.8	2379	2	Q6U9W8_PRRSV	Q6u9w8_porcine rep	630	30	81.1	227	2	Q54078_RIOSH	Q54078 rhodobacter
558	31	83.8	2606	2	Q36414_9GAMA	Q36414_alcelaphine	631	30	81.1	229	2	Q6JN61_XANCA	Q6jn61 xanthomonas
559	31	83.8	3010	2	Q9QIX3_9HEPC	Q9qix3_hepatitis c	632	30	81.1	230	2	Q5P6D9_AZOSE	Q5p6d9 azoarcus sp
560	31	83.8	3010	2	Q9QIX4_9HEPC	Q9qix4_hepatitis c	633	30	81.1	231	2	Q4XTR4_CABEL	Q4xtr4 caenorhabdi
561	31	83.8	4981	2	Q9SOR3_STRAW	Q9sor3_streptomyce	634	30	81.1	238	2	Q419Y6_GIBZE	Q419y6 gibberella
562	31	83.8	6675	2	Q4KCD8_PSEF5	Q4kcd8_pseudomonas	635	30	81.1	240	2	Q7EVB9_ORISA	Q7evb9 oryza sativ
563	30	81.1	85	2	Q4VG77_PLABE	Q4vg77_plasmodium	636	30	81.1	241	2	Q7PR77_ANOGA	Q7prt7 anopheles g
564	30	81.1	55	2	Q607C9_METCA	Q607c9_methylococc	637	30	81.1	248	1	TPIS_DEBHA	Q6bmb8 debaryomyce
565	30	81.1	73	2	Q9J2H5_9GAMA	Q9j2h5_rhesus monk	638	30	81.1	248	2	Q5N4E5_SYNP6	Q5n4e5 synechococc
566	30	81.1	75	2	Q93EZ3_SHIFL	Q93ez3_shigella fl	639	30	81.1	250	2	Q9RSP1_DEIRA	Q9rep1 deinococcoc
567	30	81.1	78	2	Q28373_HORSE	Q28373_equis cabal	640	30	81.1	251	2	Q4R9D7_MACFA	Q4r9d7 macaca fasc
568	30	81.1	84	2	Q4B295_TETNG	Q4b295_tetradoon n	641	30	81.1	254	2	Q6TZF4_9HEPC	Q6tzf4 hepatitis c
569	30	81.1	85	2	Q5Z8E4_ORISA	Q5z8e4_oryza sativ	642	30	81.1	254	2	Q6U0C3_9HEPC	Q6u0c3 hepatitis c
570	30	81.1	87	2	Q5JR97_HUMAN	Q5jr97_homo sapien	643	30	81.1	254	2	Q6U4K1_9HEPC	Q6u4k1 hepatitis c
571	30	81.1	102	2	Q4M0I3_9BURK	Q4m0i3_burkholderi	644	30	81.1	254	2	Q6U4K2_9HEPC	Q6u4k2 hepatitis c
572	30	81.1	108	2	Q6BD33_HUMAN	Q6bd33_homo sapien	645	30	81.1	254	2	Q6U4K3_9HEPC	Q6u4k3 hepatitis c
573	30	81.1	109	2	Q8WY77_HUMAN	Q8wy77_homo sapien	646	30	81.1	254	2	Q6U4K4_9HEPC	Q6u4k4 hepatitis c
574	30	81.1	115	2	Q9P7Y2_SCHPO	Q9p7y2_schizosacch	647	30	81.1	254	2	Q6U4K5_9HEPC	Q6u4k5 hepatitis c
575	30	81.1	121	2	Q66VC9_LEIDO	Q66vc9_leishmania	648	30	81.1	254	2	Q6U4K6_9HEPC	Q6u4k6 hepatitis c
576	30	81.1	121	2	Q8C3V1_MOUSE	Q8c3v1_mus musculu	649	30	81.1	254	2	Q6U4K7_9HEPC	Q6u4k7 hepatitis c
577	30	81.1	123	2	Q65178_ASP	Q65178_african swi	650	30	81.1	254	2	Q6U4K8_9HEPC	Q6u4k8 hepatitis c
578	30	81.1	128	2	Q8N8A0_HUMAN	Q8n8a0_homo sapien	651	30	81.1	254	2	Q6U4K9_9HEPC	Q6u4k9 hepatitis c
579	30	81.1	129	2	Q4UK21_RICFE	Q4uk21_rickettsia	652	30	81.1	254	2	Q6U4L0_9HEPC	Q6u4l0 hepatitis c
580	30	81.1	131	2	Q64DU2_9ARCH	Q64du2_uncultured	653	30	81.1	254	2	Q6U4L1_9HEPC	Q6u4l1 hepatitis c
581	30	81.1	132	2	Q5N5H8_SYNP6	Q5n5h8_synechococc	654	30	81.1	254	2	Q6U4L3_9HEPC	Q6u4l3 hepatitis c
582	30	81.1	134	2	Q5P3P8_AZOSE	Q5p3p8_azoarcus sp	655	30	81.1	254	2	Q6U4L4_9HEPC	Q6u4l4 hepatitis c
583	30	81.1	136	2	Q9Y9X1_AERPE	Q9y9x1_aeropyrum p	656	30	81.1	254	2	Q6U4L5_9HEPC	Q6u4l5 hepatitis c
584	30	81.1	139	2	Q4TG84_TETNG	Q4tg84_tetradoon n	657	30	81.1	254	2	Q6U4L6_9HEPC	Q6u4l6 hepatitis c
585	30	81.1	148	2	Q5CDG1_CRYHO	Q5cdg1_cryptospori	658	30	81.1	254	2	Q6U4L7_9HEPC	Q6u4l7 hepatitis c
586	30	81.1	149	2	Q5VLZ0_HALMA	Q5vlz0_haloarcula	659	30	81.1	254	2	Q6U4L8_9HEPC	Q6u4l8 hepatitis c
587	30	81.1	152	2	Q55YB3_CRYNE	Q55yb3_cryptococcu	660	30	81.1	254	2	Q6U4L9_9HEPC	Q6u4l9 hepatitis c
588	30	81.1	152	2	Q5KLM7_CRYNE	Q5klm7_cryptococcu	661	30	81.1	254	2	Q6U4M0_9HEPC	Q6u4m0 hepatitis c
589	30	81.1	152	2	Q7TFM0_RHCW6	Q7tfm0_rhesus cyto	662	30	81.1	254	2	Q6U4M1_9HEPC	Q6u4m1 hepatitis c
590	30	81.1	153	2	Q7FQO3_ORISA	Q7fqo3_oryza sativ	663	30	81.1	254	2	Q6U4M3_9HEPC	Q6u4m3 hepatitis c
591	30	81.1	156	2	Q9W8S9_9BETA	Q9w8s9_human herpe	664	30	81.1	254	2	Q6U4M4_9HEPC	Q6u4m4 hepatitis c
592	30	81.1	157	2	Q84XQ6_BRARP	Q84xq6_brassica ra	665	30	81.1	254	2	Q6U4M5_9HEPC	Q6u4m5 hepatitis c
593	30	81.1	160	2	Q4TXJ3_TETNG	Q4txj3_tetradoon n	666	30	81.1	254	2	Q6U4M6_9HEPC	Q6u4m6 hepatitis c
594	30	81.1	163	2	Q7TWV7_MYCBO	Q7twv7_mycobacteri	667	30	81.1	254	2	Q6U4M7_9HEPC	Q6u4m7 hepatitis c
595	30	81.1	163	2	Q53351_MYCTU	Q53351_mycobacteri	668	30	81.1	254	2	Q6U4M8_9HEPC	Q6u4m8 hepatitis c
596	30	81.1	166	2	Q51KZ7_9HIV1	Q51kz7_human immun	669	30	81.1	254	2	Q6U4M9_9HEPC	Q6u4m9 hepatitis c
597	30	81.1	168	2	Q7EZX8_ORISA	Q7exx8_oryza sativ	670	30	81.1	254	2	Q6U4N0_9HEPC	Q6u4n0 hepatitis c
598	30	81.1	173	2	Q89Y86_BRAJA	Q89y86_bradyrhizob	671	30	81.1	254	2	Q6U4N1_9HEPC	Q6u4n1 hepatitis c
599	30	81.1	174	2	Q6EYF1_ORISA	Q6epy1_oryza sativ	672	30	81.1	254	2	Q6U4N2_9HEPC	Q6u4n2 hepatitis c
600	30	81.1	175	2	Q5Z545_ORISA	Q5z545_oryza sativ	673	30	81.1	254	2	Q6U4N3_9HEPC	Q6u4n3 hepatitis c
601	30	81.1	179	2	Q27803_METTH	Q27803_methanobact	674	30	81.1	254	2	Q6U4N4_9HEPC	Q6u4n4 hepatitis c
602	30	81.1	181	2	Q4TSG3_9SPHN	Q4tsg3_erythrobact	675	30	81.1	254	2	Q6U4N5_9HEPC	Q6u4n5 hepatitis c
603	30	81.1	188	2	Q5TUP8_ANOGA	Q5tup8_anopheles g	676	30	81.1	254	2	Q6U4N6_9HEPC	Q6u4n6 hepatitis c
604	30	81.1	190	2	Q6D2P4_ERWCT	Q6d2p4_erwinia car	677	30	81.1	254	2	Q6U4N7_9HEPC	Q6u4n7 hepatitis c
605	30	81.1	191	2	Q6U2T2_9HEPC	Q6u2t2_hepatitis c	678	30	81.1	254	2	Q6U4N8_9HEPC	Q6u4n8 hepatitis c
606	30	81.1	191	2	Q6U2T8_9HEPC	Q6u2t8_hepatitis c	679	30	81.1	254	2	Q6U4N9_9HEPC	Q6u4n9 hepatitis c
607	30	81.1	194	2	Q9A6M7_CAUCR	Q9a6m7_caulobacter	680	30	81.1	254	2	Q6U4P0_9HEPC	Q6u4p0 hepatitis c
608	30	81.1	197	2	Q92XP7_RHIME	Q92xp7_rhizobium m	681	30	81.1	254	2	Q6U4P1_9HEPC	Q6u4p1 hepatitis c
609	30	81.1	198	2	Q96S77_HUMAN	Q96s77_homo sapien	682	30	81.1	254	2	Q6U4P2_9HEPC	Q6u4p2 hepatitis c
610	30	81.1	199	2	Q7ZP17_9HIV1	Q7zpl7_human immun	683	30	81.1	254	2	Q6U4P3_9HEPC	Q6u4p3 hepatitis c
611	30	81.1	200	2	Q8GSG9_WHEAT	Q8gsq9_triticum ae	684	30	81.1	254	2	Q6U4P5_9HEPC	Q6u4p5 hepatitis c
612	30	81.1	201	2	Q57P18_SALCH	Q57p18_salmonella	685	30	81.1	254	2	Q6U4P6_9HEPC	Q6u4p6 hepatitis c
613	30	81.1	202	2	Q6K3R4_ORISA	Q6k3r4_oryza sativ	686	30	81.1	254	2	Q6U4P7_9HEPC	Q6u4p7 hepatitis c
614	30	81.1	202	2	Q7V7H9_PROMM	Q7v7h9_prochloroco	687	30	81.1	254	2	Q6U4Q4_9HEPC	Q6u4q4 hepatitis c
615	30	81.1	202	2	Q70QX2_9HIV1	Q70qx2_human immun	688	30	81.1	254	2	Q6U4R2_9HEPC	Q6u4r2 hepatitis c

689	30	81.1	254	2	Q6U594_9HEPC	Q6U594	hepatitis c	762	30	81.1	329	2	Q76BK6_PAROL	Q76bk6	paralichthy
690	30	81.1	254	2	Q6U5A3_9HEPC	Q6U5A3	hepatitis c	763	30	81.1	333	2	Q8KQV3_VIBCH	Q8kqv3	vibrio chol
691	30	81.1	254	2	Q6U5A4_9HEPC	Q6U5A4	hepatitis c	764	30	81.1	335	2	Q9TM68_ALECA	Q9tm68	alexandrium
692	30	81.1	254	2	Q6U5A5_9HEPC	Q6U5A5	hepatitis c	765	30	81.1	338	2	Q6ZGQ0_ORISA	Q6zgg0	oryza sativ
693	30	81.1	254	2	Q6U5A7_9HEPC	Q6U5A7	hepatitis c	766	30	81.1	343	2	Q9TM72_PROMC	Q9tm72	prorocentrum
694	30	81.1	254	2	Q6U5B0_9HEPC	Q6U5B0	hepatitis c	767	30	81.1	345	2	Q5SDQ5_DICTDI	Q5sdq5	dictyosteli
695	30	81.1	254	2	Q6U5B2_9HEPC	Q6U5B2	hepatitis c	768	30	81.1	346	2	Q82MX3_STRAW	Q82mx3	strepomyce
696	30	81.1	254	2	Q6U5B4_9HEPC	Q6U5B4	hepatitis c	769	30	81.1	346	2	Q7NET1_GLOVI	Q7net1	gloeobacter
697	30	81.1	254	2	Q6U5B5_9HEPC	Q6U5B5	hepatitis c	770	30	81.1	347	2	Q7NET6_ALETA	Q7net6	alexandrium
698	30	81.1	254	2	Q6U5B6_9HEPC	Q6U5B6	hepatitis c	771	30	81.1	347	2	Q8B3W0_9GAMA	Q8b3w0	porcine lym
699	30	81.1	254	2	Q6U5B7_9HEPC	Q6U5B7	hepatitis c	772	30	81.1	348	2	Q5DFV0_SCHJA	Q5dfv0	schistosoma
700	30	81.1	254	2	Q6U5B8_9HEPC	Q6U5B8	hepatitis c	773	30	81.1	350	2	Q5BBY4_EMENI	Q5bby4	aspergillus
701	30	81.1	254	2	Q6U5C0_9HEPC	Q6U5C0	hepatitis c	774	30	81.1	350	2	Q7XHJ4_WHEAT	Q7xhj4	tritricum ae
702	30	81.1	254	2	Q6U5C4_9HEPC	Q6U5C4	hepatitis c	775	30	81.1	352	2	Q9VWY7_DROME	Q9vwy7	drosofila
703	30	81.1	254	2	Q6U5C7_9HEPC	Q6U5C7	hepatitis c	776	30	81.1	354	2	Q5JJC2_PYROCO	Q5jjc2	pyrococcus
704	30	81.1	254	2	Q6U5D3_9HEPC	Q6U5D3	hepatitis c	777	30	81.1	357	2	Q8KX2_PRORE	Q8kx2	providencia
705	30	81.1	254	2	Q6U5D4_9HEPC	Q6U5D4	hepatitis c	778	30	81.1	363	2	Q8Y286_ANASP	Q8y286	anabaena sp
706	30	81.1	254	2	Q6U5D5_9HEPC	Q6U5D5	hepatitis c	779	30	81.1	365	2	Q9SZ84_ARATH	Q9szf4	arabidopsis
707	30	81.1	254	2	Q6U5D6_9HEPC	Q6U5D6	hepatitis c	780	30	81.1	367	1	CD5R2_HUMAN	Q13319	homo sapien
708	30	81.1	256	2	Q657H9_ORISA	Q657H9	oryza sativ	781	30	81.1	367	2	Q4ZFW6_HUMAN	Q4zfw6	homo sapien
709	30	81.1	257	2	Q75GB1_ORISA	Q75gb1	oryza sativ	782	30	81.1	367	2	Q5XKD4_HUMAN	Q5xkd4	homo sapien
710	30	81.1	257	2	Q4TFG9_TETNG	Q4tf99	tetradodon n	783	30	81.1	367	2	Q8LFA7_ARATH	Q8lfa7	arabidopsis
711	30	81.1	259	2	Q84CV7_9BACT	Q84cv7	uncultured	784	30	81.1	368	2	Q8GM52_ARATH	Q8gm52	arabidopsis
712	30	81.1	260	2	Q6GIL7_HUMAN	Q6gil7	homo sapien	785	30	81.1	368	2	Q6EJ98_ARATH	Q6ej98	arabidopsis
713	30	81.1	260	2	Q7PYB1_ANOGA	Q7pyb1	homo sapien	786	30	81.1	368	2	Q5K481_SILDC	Q5k481	silene dici
714	30	81.1	262	1	SP1B_HUMAN	Q01892	homo sapien	787	30	81.1	368	2	Q744W8_MYCPA	Q744w8	mycobacteri
715	30	81.1	262	2	Q5NVX0_9ARCH	Q5nvx0	uncultured	788	30	81.1	370	2	Q5K480_SILDC	Q5k480	silene cucu
716	30	81.1	262	2	Q523C2_NOCFA	Q523c2	nocardia fa	789	30	81.1	370	2	Q5K482_SILDC	Q5k482	silene dici
717	30	81.1	262	2	Q9JMW6_RAT	Q9jmw6	rattus norv	790	30	81.1	370	2	Q8YL37_ANASP	Q8yl37	anabaena sp
718	30	81.1	262	2	Q6UDZ4_HCMV	Q6udz4	human cytom	791	30	81.1	371	2	Q8LGL6_SILLA	Q8lgl6	silene lati
719	30	81.1	264	2	Q96C55_HUMAN	Q96c55	homo sapien	792	30	81.1	372	2	Q5K484_SILDI	Q5k484	silene dioi
720	30	81.1	264	2	Q6NW31_HUMAN	Q6nw31	homo sapien	793	30	81.1	372	2	Q8LGL1_SILDI	Q8lgl1	silene dioi
721	30	81.1	266	2	Q52MA4_XENTR	Q52ma4	xenopus tro	794	30	81.1	372	2	Q8LGL2_9CARY	Q8lgl2	lychnis fio
722	30	81.1	267	2	Q6ZMK6_HUMAN	Q6zmk6	homo sapien	795	30	81.1	373	2	Q8LGL7_SILCU	Q8lgl7	silene lati
723	30	81.1	268	2	Q8T0C9_DROME	Q8t0c9	drosofila	796	30	81.1	373	2	Q8LGL4_SILCU	Q8lgl4	silene cucu
724	30	81.1	268	2	Q6N9Z3_RHOPA	Q6n9z3	rhodospseudo	797	30	81.1	373	2	Q8RXB3_SILCO	Q8rxb3	silene coni
725	30	81.1	269	2	Q4W9L6_ASPPU	Q4w9l6	aspergillus	798	30	81.1	373	2	Q4UPL5_XANCP	Q4upl5	xanthomonas
726	30	81.1	269	2	Q54SR6_DICTDI	Q54sr6	dictyosteli	799	30	81.1	373	2	Q8P434_XANCP	Q8p434	xanthomonas
727	30	81.1	269	2	Q6H714_ORISA	Q6h714	oryza sativ	800	30	81.1	376	2	Q9U5J1_9TRYP	Q9u5j1	trypanosoma
728	30	81.1	271	2	Q915J4_PSEAB	Q915j4	pseudomonas	801	30	81.1	376	2	Q8VWN5_SILLA	Q8vwn5	silene lati
729	30	81.1	272	2	Q8RVM1_MUSPK	Q8rvw1	musa acumin	802	30	81.1	376	2	Q72E67_DESVH	Q72e67	desulfovibr
730	30	81.1	272	2	Q8RRF6_PSEPK	Q8rrf6	pseudomonas	803	30	81.1	380	2	Q4XYH3_PLACH	Q4xyh3	plasmodium
731	30	81.1	272	2	Q99D13_BHV4	Q99d13	bovine herp	804	30	81.1	384	2	Q67M96_ORISA	Q67m96	oryza sativ
732	30	81.1	275	2	Q62364_CABEL	Q62364	caenorhabdi	805	30	81.1	386	2	Q7WYR3_RHILV	Q7wyr3	rhizobium l
733	30	81.1	278	1	MA32_MOUSE	Q35658	mus musculus	806	30	81.1	387	2	Q4P8Y3_USITAGO	Q4p8y3	ustilago ma
734	30	81.1	278	2	Q8H769_WHEAT	Q8h769	triticum ae	807	30	81.1	391	2	Q97553_CERAE	Q97553	cercopithe
735	30	81.1	279	2	Q4OCN2_LEIMA	Q4ocn2	leishmania	808	30	81.1	392	2	Q5E9Q1_BOVIN	Q5e9q1	bos taurus
736	30	81.1	279	2	Q9RWR7_DEIRA	Q9rwr7	deinococcus	809	30	81.1	392	2	Q6GR57_XENLA	Q6gr57	xenopus lae
737	30	81.1	279	2	Q8RSL1_MOUSE	Q8rsl1	mus musculus	810	30	81.1	393	2	Q7UXV3_RHOBA	Q7uxv3	rhodopirell
738	30	81.1	279	2	Q61R85_RAT	Q61r85	rattus norv	811	30	81.1	395	2	Q4H9L0_9DEIO	Q4h9l0	deinococcus
739	30	81.1	281	2	Q75AG0_ASHGO	Q75ago	ashbya gos	812	30	81.1	395	2	Q8YP52_ANASP	Q8yp52	anabaena sp
740	30	81.1	281	2	Q5DPB3_SCHJA	Q5dpb3	schistosoma	813	30	81.1	396	2	Q4WJ66_ASPPU	Q4wj66	aspergillus
741	30	81.1	281	2	Q9VIX3_DROME	Q9vix3	drosofila	814	30	81.1	397	2	Q8R0W1_MOUSE	Q8r0w1	mus musculu
742	30	81.1	281	2	Q45765_CABEL	Q45765	caenorhabdi	815	30	81.1	398	2	Q6K6D7_ORISA	Q6k6d7	oryza sativ
743	30	81.1	283	2	Q52W85_9DINO	Q52w85	akashiwos	816	30	81.1	402	2	Q85717_SPRCU	Q85717	strepomyce
744	30	81.1	294	2	Q69PC5_ORISA	Q69pc5	oryza sativ	817	30	81.1	402	2	Q829Q6_STRAW	Q829q6	strepomyce
745	30	81.1	294	2	Q6K337_ORISA	Q6k337	oryza sativ	818	30	81.1	403	2	Q4RV84_TETNG	Q4rv84	tetradodon n
746	30	81.1	295	2	Q5FPH3_GLUOX	Q5fph3	gluconobact	819	30	81.1	405	2	Q8LDF1_ARATH	Q8ldf1	arabidopsis
747	30	81.1	295	2	Q732Q4_MYCPA	Q732q4	mycobacteri	820	30	81.1	405	2	Q94BZ6_ARATH	Q94bz6	arabidopsis
748	30	81.1	296	2	Q9FVR6_ARATH	Q9fvr6	arabidopsis	821	30	81.1	408	2	Q7X692_ORISA	Q7x692	oryza sativ
749	30	81.1	299	2	Q9M4PT_BRANA	Q9m4f1	brassica na	822	30	81.1	411	2	Q5N397_SYNPE	Q5n397	synchococc
750	30	81.1	300	2	Q4ZLM2_PSESY	Q4zlm2	pseudomonas	823	30	81.1	413	2	Q5YUF8_NOCFA	Q5yuf8	nocardia fa
751	30	81.1	302	2	Q9CSV8_MOUSE	Q9cev8	mus musculus	824	30	81.1	413	2	Q7SZU5_XENLA	Q7szu5	xenopus lae
752	30	81.1	302	2	Q949E0_ORISA	Q949e0	oryza sativ	825	30	81.1	413	2	Q8Q3W6_XENLA	Q8q3w6	xenopus lae
753	30	81.1	303	2	Q4LGG1_9BURK	Q4lgg1	burkholderi	826	30	81.1	413	2	Q90XS5_XENLA	Q90xs5	xenopus lae
754	30	81.1	302	2	Q8S9V8_SHEON	Q8sv8	shevanelia	827	30	81.1	419	2	Q9RD81_STRCO	Q9rd81	strepomyce
755	30	81.1	304	1	NIF51_ARATH	Q9sv84	arabidopsis	828	30	81.1	419	2	Q8RIN4_STRCO	Q8rit4	strepomyce
756	30	81.1	311	2	Q58CX8_BOVIN	Q58cx8	bos taurus	829	30	81.1	422	2	Q8VWN3_SILLA	Q8vwn3	silene lati
757	30	81.1	312	2	Q9LHF0_ARATH	Q9lhf0	arabidopsis	830	30	81.1	425	2	Q4V626_DROME	Q4v626	drosofila
758	30	81.1	314	2	Q82C21_STRAW	Q82c21	strepomyce	831	30	81.1	425	2	Q8VWN2_SILLA	Q8vwn2	silene lati
759	30	81.1	322	2	Q6R960_9RHIZ	Q6r960	hyphomicrob	832	30	81.1	425	2	Q9LLR5_JUNEF	Q9llr5	silene effu
760	30	81.1	324	2	Q7RZV6_NEUCR	Q7rvz6	neurospora	833	30	81.1	426	2	Q8VWN4_SILLA	Q8vwn4	silene lati
761	30	81.1	329	2	Q8GHY1_PSBRE	Q8ghy1	pseudomonas	834	30	81.1	430	2	Q67N20_SYMTH	Q67n20	symbiobacte

835	30	81.1	431	2	Q8U119_AGR75	Q8U119_agrobacteri	908	30	81.1	545	2	Q4P5G9_USTMA	Q4P5G9_ustilago ma
836	30	81.1	435	2	Q7D1D9_AGR75	Q7D1D9_agrobacteri	909	30	81.1	549	2	Q8GUJ8_ARATH	Q8GUJ8_arabidopsis
837	30	81.1	438	2	Q5OYN2_ENTHI	Q5OYN2_entamoeba h	910	30	81.1	553	2	Q6IG04_RAT	Q6IG04_rattus norv
838	30	81.1	440	2	Q8MRX2_DROME	Q8MRX2_drosophila	911	30	81.1	558	2	Q61J10_CAEBR	Q61J10_caenorhabdi
839	30	81.1	440	2	Q9V871_DROME	Q9V871_drosophila	912	30	81.1	559	2	Q4PDA0_USTMA	Q4PDA0_ustilago ma
840	30	81.1	441	2	Q5DFB1_SCHJA	Q5DFB1_schistosoma	913	30	81.1	560	2	Q7NJ14_GLOVI	Q7NJ14_gloeobacter
841	30	81.1	441	2	Q9VAL8_DROME	Q9VAL8_drosophila	914	30	81.1	561	2	Q4WAY7_ASFFU	Q4WAY7_aspergillus
842	30	81.1	442	2	Q88HG2_PSPPK	Q88HG2_pseudomonas	915	30	81.1	562	2	Q6GPL0_XENLA	Q6GPL0_xenopus lae
843	30	81.1	443	2	Q80T99_MOUSE	Q80T99_mus musculus	916	30	81.1	565	2	Q59ZZ2_CANAL	Q59ZZ2_candida alb
844	30	81.1	444	2	Q52596_BRAJA	Q52596_bradyrhizob	917	30	81.1	565	2	Q6CDR6_YARLI	Q6CDR6_yarrowia li
845	30	81.1	447	2	Q5E914_BOVIN	Q5E914_bos taurus	918	30	81.1	566	2	Q8CAC0_MOUSE	Q8CAC0_mus musculus
846	30	81.1	447	2	Q8BOD8_9HEPC	Q8BOD8_hepatitis c	919	30	81.1	567	2	Q5A044_CANAL	Q5A044_candida alb
847	30	81.1	447	2	Q8BOD9_9HEPC	Q8BOD9_hepatitis c	920	30	81.1	567	2	Q51BC8_ENTHI	Q51BC8_entamoeba h
848	30	81.1	449	2	Q6N011_RHOPA	Q6N011_rhodospheudo	921	30	81.1	567	2	Q9LPW5_ARATH	Q9LPW5_arabidopsis
849	30	81.1	453	2	Q5CHG1_CRYHO	Q5CHG1_cryptospori	922	30	81.1	568	2	Q6ZN96_HUMAN	Q6ZN96_homo sapien
850	30	81.1	455	2	Q69MP5_ORYSA	Q69MP5_oryza sativ	923	30	81.1	570	2	Q5IIP0_MAGGR	Q5IIP0_magnaporthe
851	30	81.1	457	2	Q60WL5_CAEBR	Q60WL5_caenorhabdi	924	30	81.1	571	2	Q6Q138_BOVIN	Q6Q138_bos taurus
852	30	81.1	458	2	Q9SY19_ARATH	Q9SY19_arabidopsis	925	30	81.1	579	2	Q8QZS3_MOUSE	Q8QZS3_mus musculus
853	30	81.1	461	2	Q4TB03_TETNG	Q4TB03_tetradodon n	926	30	81.1	579	2	Q76JQ2_RAT	Q76JQ2_rattus norv
854	30	81.1	462	2	Q72735_HUMAN	Q72735_homo sapien	927	30	81.1	579	2	Q6SWP7_HCMV	Q6SWP7_human cytom
855	30	81.1	462	2	Q68A51_RALSO	Q68A51_ralstonia s	928	30	81.1	580	1	VIE2_HCMVA	VIE2_human cytom
856	30	81.1	462	2	Q5LNV7_SILPO	Q5LNV7_silicibacte	929	30	81.1	580	2	Q5IU02_MAGGR	Q5IU02_magnaporthe
857	30	81.1	462	2	Q8XYF8_RALSO	Q8XYF8_ralstonia s	930	30	81.1	580	2	Q6SW29_HCMV	Q6SW29_human cytom
858	30	81.1	463	2	Q51BD2_ENTHI	Q51BD2_entamoeba h	931	30	81.1	580	2	Q6SWJ2_HCMV	Q6SWJ2_human cytom
859	30	81.1	463	2	Q6N6F9_RHOPA	Q6N6F9_rhodospheudo	932	30	81.1	580	2	Q6SML9_HCMV	Q6SML9_human cytom
860	30	81.1	464	2	Q9KHK2_9ACTO	Q9KHK2_streptomyce	933	30	81.1	580	2	Q6SMSE_HCMV	Q6SMSE_human cytom
861	30	81.1	466	2	Q5IAU5_ENTHI	Q5IAU5_entamoeba h	934	30	81.1	580	2	Q6SWY2_HCMV	Q6SWY2_human cytom
862	30	81.1	469	2	Q5EIT5_LISSE	Q5EIT5_listeria se	935	30	81.1	580	2	Q6SX11_HCMV	Q6SX11_human cytom
863	30	81.1	469	2	Q8VMW8_LISSE	Q8VMW8_listeria se	936	30	81.1	580	2	Q6SX40_HCMV	Q6SX40_human cytom
864	30	81.1	469	2	Q9A3Y2_CAUCR	Q9A3Y2_caulobacter	937	30	81.1	580	2	Q6SWV3_HCMV	Q6SWV3_human cytom
865	30	81.1	470	2	Q6WMG4_DBDBA	Q6WMG4_bdellovibri	938	30	81.1	582	2	Q98683_9BETA	Q98683_simian cyto
866	30	81.1	475	2	Q6PKN2_9STRA	Q6PKN2_nitzschia c	939	30	81.1	586	2	Q8NA33_HUMAN	Q8NA33_homo sapien
867	30	81.1	476	2	Q9V7E4_DROME	Q9V7E4_drosophila	940	30	81.1	595	2	Q73S88_MTCPA	Q73S88_mycobacteri
868	30	81.1	476	2	Q9LLA9_LYCES	Q9LLA9_lycopersico	941	30	81.1	595	2	Q82E16_STRAW	Q82E16_streptomyce
869	30	81.1	477	1	Q9ARH8_LYCES	Q9ARH8_lycopersico	942	30	81.1	598	2	Q8QRY7_9BETA	Q8QRY7_pongine her
870	30	81.1	477	1	MAZ_HUMAN	P56270_homo sapien	943	30	81.1	600	2	Q6AZT2_XENLA	Q6AZT2_xenopus lae
871	30	81.1	477	1	MAZ_MOUSE	P56571_mus musculus	944	30	81.1	603	2	Q74DH2_GOSL	Q74DH2_geobacter s
872	30	81.1	477	2	Q8W939_9ASTE	Q8W939_phyllonoma	945	30	81.1	603	2	Q9PT11_XENLA	Q9PT11_xenopus lae
873	30	81.1	478	2	Q41Y77_GIBZEL	Q41Y77_gibberella	946	30	81.1	607	2	Q4Q7G9_LEIMA	Q4Q7G9_leishmania
874	30	81.1	478	2	Q4WJN9_ASFFU	Q4WJN9_aspergillus	947	30	81.1	609	2	Q94BT6_ARATH	Q94BT6_arabidopsis
875	30	81.1	479	2	Q9XG18_LYCES	Q9XG18_lycopersico	948	30	81.1	609	2	Q94LDF6_ARATH	Q94LDF6_arabidopsis
876	30	81.1	480	2	Q9ARH7_CAPAN	Q9ARH7_capsicum an	949	30	81.1	610	2	Q7QTH7_GIALA	Q7QTH7_giardia lam
877	30	81.1	480	2	Q39443_CAPAN	Q39443_capsicum an	950	30	81.1	612	2	Q8RUM5_ORYSA	Q8RUM5_oryza sativ
878	30	81.1	481	2	Q8CXN8_OCEIH	Q8CXN8_oceanobacil	951	30	81.1	612	2	Q941S9_ORYSA	Q941S9_oryza sativ
879	30	81.1	482	2	Q59L98_CANAL	Q59L98_candida alb	952	30	81.1	615	1	TUP11_SCHPO	TUP11_schizosacch
880	30	81.1	488	2	Q89WD2_BRAJA	Q89WD2_bradyrhizob	953	30	81.1	615	2	Q75DI6_ASHGO	Q75DI6_ashbya gos
881	30	81.1	488	2	Q6UAQ1_TETNG	Q6UAQ1_tetradodon n	954	30	81.1	615	2	Q4UEI5_MESECR	Q4UEI5_mesembryant
882	30	81.1	490	2	Q82B66_STRAW	Q82B66_streptomyce	955	30	81.1	621	2	Q4RM99_TBTNG	Q4RM99_tetradodon n
883	30	81.1	491	2	Q5NKH1_AZOSE	Q5NKH1_azearcus sp	956	30	81.1	622	2	Q617J8_CAEBR	Q617J8_caenorhabdi
884	30	81.1	492	1	CP26A_CHICK	Q9PUB4_gallus gall	957	30	81.1	623	2	Q8BUU5_DBDBA	Q8BUU5_debaryomyce
885	30	81.1	493	2	Q8NPN7_HUMAN	Q8NPN7_homo sapien	958	30	81.1	623	2	Q8WZA9_HUMAN	Q8WZA9_homo sapien
886	30	81.1	496	2	Q93YF8_9SOLA	Q93YF8_nicotiana a	959	30	81.1	624	1	SIR_SYNP7	SIR_synethococci
887	30	81.1	500	2	Q4WC97_ASFFU	Q4WC97_aspergillus	960	30	81.1	626	2	Q6Q079_CAEBR	Q6Q079_caenorhabdi
888	30	81.1	500	2	Q24348_SORBI	Q24348_sorghum bic	961	30	81.1	628	2	Q41R92_GIBZEL	Q41R92_gibberella
889	30	81.1	502	1	AKA80_RAT	Q35817_rattus norv	962	30	81.1	630	2	Q528K3_ORYSA	Q528K3_oryza sativ
890	30	81.1	502	2	Q4FJW3_BOVIN	Q4FJW3_bos taurus	963	30	81.1	630	2	Q8PMC2_XANAC	Q8PMC2_xanthomonas
891	30	81.1	502	2	Q7UE37_RHOBA	Q7UE37_rhodopirell	964	30	81.1	633	2	Q5N202_SYNP6	Q5N202_synethococci
892	30	81.1	503	2	Q4WJ30_ASFFU	Q4WJ30_aspergillus	965	30	81.1	634	1	PHAR2_HUMAN	PHAR2_homo sapien
893	30	81.1	506	2	Q413L5_GIBZEL	Q413L5_gibberella	966	30	81.1	634	2	Q5TFA0_HUMAN	Q5TFA0_homo sapien
894	30	81.1	509	2	Q41BZ6_GIBZEL	Q41BZ6_gibberella	967	30	81.1	635	2	Q5LS83_SILPO	Q5LS83_silicibacte
895	30	81.1	509	2	Q60979_LEIMA	Q60979_leishmania	968	30	81.1	637	2	Q510G4_RAT	Q510G4_rattus norv
896	30	81.1	510	1	CX366_CHICK	P29415_gallus gall	969	30	81.1	641	2	Q17740_CABEL	Q17740_caenorhabdi
897	30	81.1	517	2	Q7S9S1_NEUCR	Q7S9S1_neurospora	970	30	81.1	641	2	Q7ZZW2_PAROL	Q7ZZW2_paralichthy
898	30	81.1	519	2	Q41JF2_GIBZEL	Q41JF2_gibberella	971	30	81.1	642	2	Q6BQJ8_DBDBA	Q6BQJ8_debaryomyce
899	30	81.1	522	2	Q6PAZ2_XENLA	Q6PAZ2_xenopus lae	972	30	81.1	642	2	Q4YN59_PLABE	Q4YN59_plasmodium
900	30	81.1	524	2	Q5AWB5_EMENI	Q5AWB5_aspergillus	973	30	81.1	643	2	Q7TYP4_MYCBO	Q7TYP4_mycobacteri
901	30	81.1	527	2	Q75191_HUMAN	Q75191_homo sapien	974	30	81.1	643	2	P71750_MYCTU	P71750_mycobacteri
902	30	81.1	532	2	Q71EE9_9HIV1	Q71EE9_human immun	975	30	81.1	644	2	Q8DL77_SYNEL	Q8DL77_synethococci
903	30	81.1	534	2	Q8U3K6_PYRWM	Q8U3K6_pyrococcus	976	30	81.1	645	2	Q4WQ81_ASFFU	Q4WQ81_aspergillus
904	30	81.1	536	2	Q7V997_PROCHLO	Q7V997_prochloroco	977	30	81.1	645	2	Q6PJ32_HUMAN	Q6PJ32_homo sapien
905	30	81.1	541	1	GATB_YEAST	P33893_saccharomyc	978	30	81.1	647	2	Q5Y253_CRYGA	Q5Y253_cryptococc
906	30	81.1	542	1	Q9594_CAEEL	Q9594_caenorhabdi	979	30	81.1	650	2	Q8YX67_ANAPR	Q8YX67_anabaena sp
907	30	81.1	544	2	Q7M7T5_WOLSU	Q7M7T5_wolinnella s	980	30	81.1	650	2	Q58EJ8_BRARE	Q58EJ8_brachydanio

```

981 30 81.1 652 2 Q7R4B0 GIALA
982 30 81.1 652 2 Q5N4E9 synchococc
983 30 81.1 653 2 Q5G9Y5 crassostrea
984 30 81.1 653 2 Q6GNY5 xenopus lae
985 30 81.1 659 2 Q5AXZ0 aspergillus
986 30 81.1 660 2 Q4PGD1 USTMA
987 30 81.1 664 2 Q4GZ71 9TRYP
988 30 81.1 666 2 Q5FSZ2 GUJUX
989 30 81.1 670 2 Q59385 cyprinus ca
990 30 81.1 672 2 Q613K3 CAEBR
991 30 81.1 695 2 Q8ZUD2 PYRAE
992 30 81.1 695 2 Q7SA97 NEUCR
993 30 81.1 697 1 YAYR RHISN
994 30 81.1 697 2 QRQB0 RHIFR
995 30 81.1 697 2 Q7YZ9 BRARE
996 30 81.1 703 1 CASC3 HUMAN
997 30 81.1 705 2 Q5SUT6 cryptococcu
998 30 81.1 705 2 Q5KHK2 CRYNE
999 30 81.1 708 2 Q8NLG5 HUMAN
1000 30 81.1 708 2 Q57ZF9 9TRYP

```

ALIGNMENTS

```

RESULT 1
Q5MZB3 SYN6 PRELIMINARY; PRT; 221 AA.
AC Q5MZB3,
DT 01-FEB-2005 (TREMBLrel. 29, Created)
DT 01-FEB-2005 (TREMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TREMBLrel. 29, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=svc2417.d;
OS Synchococcus sp. (strain PCC 6301) (Anacystis nidulans).
OC Bacteria; Cyanobacteria; Chroococcales; Synchococcus.
CX NCBI_TaxID=269084;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PCC6301;
RA Sugita M.;
RT "Complete genome structure of the unicellular cyanobacterium Anacystis
RT nidulans 6301 (Synchococcus sp. PCC6301)".
RL Submitted (DEC-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP008231; BAD80607.1; -; Genomic_DNA.
DR InterPro; IPR005829; Sug_transporter.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; UNKNOWN_1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 221 AA; 24058 MW; FFA89192DCEEB750 CRC64;

```

```

Query Match 100.0%; Score 37; DB 2; Length 221;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAPPVP 7
Db 122 ILAPPVP 128

```

```

RESULT 2
Q9LGV4 ORYSA PRELIMINARY; PRT; 370 AA.
AC Q9LGV4;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Hypothetical protein P0706805.32.
GN Name=P0706805.32;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
CX NCBI_TaxID=39947;

```

```

[1]
RN NUCLEOTIDE SEQUENCE.
RP Sasaki T., Matsumoto T., Yamamoto K., Sakata K., Baba T., Katayose Y.,
RA Wu J., Nimura Y., Cheng Z., Nagamura Y., Antonio B.A., Kanamori H.,
RA Hosokawa S., Masukawa M., Arikawa K., Chiden Y., Hayashi M.,
RA Okamoto M., Ando T., Aoki H., Arita K., Hamada M., Harada C.,
RA Hijishita S., Honda M., Ichikawa Y., Idonuma A., Iijima M., Ikeda M.,
RA Ikono M., Itoh S., Itoh T., Itoh Y., Iwabuchi A., Kamiya K.,
RA Karasawa M., Kasegiri S., Kikuta A., Kobayashi N., Kono I.,
RA Machita K., Maehara T., Mizuno H., Mizubayashi T., Mukai Y.,
RA Nagasaki H., Nakashima M., Nakama Y., Nakamichi Y., Nakamura M.,
RA Namiki N., Negishi M., Ohta I., Ono N., Saij S., Sakai K., Shibata M.,
RA Shimokawa T., Shomura A., Song J., Takazaki Y., Terasawa K., Taji K.,
RA Waki K., Yamagata H., Yamane H., Yoshiki S., Yoshihara R., Yukawa K.,
RA Zhong H., Iwama H., Gojibori T.;
RA Yano M., Jiang J., Gojibori T.;
RT "The genome sequence and structure of rice chromosome 1.";
RL Nature 420:312-316(2002).
DR EMBL; AP002482; BAA96629.1; -; Genomic_DNA.
DR Gramene; Q9LGV4; -.
KW Hypothetical protein.
SQ SEQUENCE 370 AA; 38964 MW; D3797D5B41FF04C9 CRC64;

Query Match 100.0%; Score 37; DB 2; Length 370;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAPPVP 7
Db 105 ILAPPVP 111

RESULT 3
Q8SXB2 DROME PRELIMINARY; PRT; 1064 AA.
ID Q8SXB2 DROME PRELIMINARY;
AC Q8SXB2;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE GH15539p.
GN Name=Ptp4E; ORFNames=CG6899;
OS Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
CX NCBI_TaxID=7227;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Berkley;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celniker S.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
DR EMBL; AY094714; AAM11067.1; -; mRNA.
DR HSP; P10586; ILAR.
DR FlyBase; FBgn0004368; CG6899.
DR FlyBase; FBgn0004368; Ptp4E.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0004725; F:protein tyrosine phosphatase activity; IEA.
DR GO; GO:0006470; P:protein amino acid dephosphorylation; IEA.
DR InterPro; IPR003962; FniII subd.
DR InterPro; IPR003961; FN III.
DR InterPro; IPR00387; TYR_phosphatase.
DR InterPro; IPR000242; Tyr_PP.
DR Pfam; PF00041; fn3; 6.
DR Pfam; PF00102; Y_phosphatase; 1.
DR PRINTS; PR00014; FNTYPSIII.
DR PRINTS; PR00700; PRTYPPHTASE.

```

DR SMART; SM00060; FN3; 6.
DR SMART; SM00194; PTPC; 1.
DR PROSITE; PS00853; FN3; 6.
DR PROSITE; PS00383; TYR PHOSPHATASE 1; 1.
DR PROSITE; PS00556; TYR PHOSPHATASE 2; 1.
DR PROSITE; PS00555; TYR PHOSPHATASE_PTP; 1.
KW Hydrolase; Repeat; Transmembrane.
SQ SEQUENCE 1064 AA; 120524 MW; 43F5DCBB8DBAEC8 CRC64;
Query Match 100.0%; Score 37; DB 2; Length 1064;
Best Local Similarity 100.0%; Pred. No. 7.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ILAPPVP 7
Db 553 ILAPPVP 559
RESULT 4
OSIRSO DROME
ID OSIRSO_DROME PRELIMINARY; PRT; 1607 AA.
AC Q8IRSO;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE CG6899-PB, isoform B.
GN Name=Pc4E; ORFNames=CG6899;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_taxid=7227;
RN NUCLEOTIDE SEQUENCE.
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA Sutton G.G., Wortman J.R., Richards S., Ashburner M., Henderson S.N.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foeller C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalil M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laske P., Lei Y., Levitsky A.C., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle B.J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
SQ Science 287:2185-2195 (2000).

RN NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426065; PubMed=12537568;
RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA George R.A., Hoskins R.A., Laverly T., Muzny D.M., Nelson C.R.,
RA Pacle B.J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Svirskaas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstock G., Scher S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT "Finishing a whole-genome shotgun: release 3 of the Drosophila
melanogaster euchromatic genome sequence.";
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426070; PubMed=12537573;
RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J.W., Svirskaas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celniker S.E.;
RT "The transposable elements of the Drosophila melanogaster euchromatin:
a genomics perspective.";
RL Genome Biol. 3:RESEARCH0084.1-RESEARCH0084.20(2002).
RN NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426069; PubMed=12537572;
RA Mira S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Willburn G.H., Prochnik S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
systematic review.";
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
RN NUCLEOTIDE SEQUENCE.
RG Berkeley Drosophila Genome Project;
RA Celniker S., Carlson J., Wan K., Pfeiffer B., Frise E., George R.,
RA Hoskins R., Stapleton M., Pacle B.J., Park S., Svirskaas R., Smith E.,
RA Yu C., Rubin G.;
RT "Drosophila melanogaster release 4 sequence.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN NUCLEOTIDE SEQUENCE.
RP Flybase;
RG Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
RL -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC EMBL; AE001432; AAN09133.1; -; Genomic_DNA.
DR HSP; P10586; ILAR.
DR Ensembl; CG6899; Drosophila melanogaster.
DR Flybase; FBgn0004368; CG6899.
DR Flybase; FBgn0004368; Ptp4E.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0004725; P:protein tyrosine phosphatase activity; IEA.
DR GO; GO:0006470; P:protein amino acid dephosphorylation; IEA.
DR InterPro; IPR003962; FNIII subd.
DR InterPro; IPR003961; FN III.
DR InterPro; IPR00387; TVR phosphatase.
DR InterPro; IPR00242; Tyr_PP.
DR Pfam; PF00041; fn3; 10.
DR Pfam; PF00102; Y_phosphatase; 1.
DR PRINTS; PR00014; FNTYPEIII.
DR PRINTS; PR00700; PRTYPHPTASE.
DR SMART; SM00060; FN3; 11.
DR SMART; SM00194; PTPC; 1.
DR PROSITE; PS00853; FN3; 11.
DR PROSITE; PS00383; TYR PHOSPHATASE 1; 1.
DR PROSITE; PS00556; TYR PHOSPHATASE 2; 1.
DR PROSITE; PS00555; TYR PHOSPHATASE_PTP; 1.
KW Hydrolase; Repeat; Transmembrane.
SQ SEQUENCE 1607 AA; 182285 MW; 6CAF554E61A27F45 CRC64;

Query Match 100.0%; Score 37; DB 2; Length 1607;
 Best Local Similarity 100.0%; Pred. No. 1.1e+03; Indels 0; Gaps 0;
 Matches 7; Conservative 0; Mismatches 0;

Qy 1 ILAPVP 7
 Db 1101 ILAPVP 1107

RESULT 5

ID Q24495 DROME PRELIMINARY; PRT; 1767 AA.
 AC Q24495
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
 DE Receptor protein tyrosine phosphatase.
 GN Name=Ptp4E;
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=94043220; PubMed=8226338;
 RA Con S.H., Hong A., Yang X., Chia W.;
 RT "Alternative splicing in a novel tyrosine phosphatase gene (DPTP4E) of
 Drosophila melanogaster generates two large receptor-like proteins
 RT which differ in their carboxyl termini.";
 RL J. Biol. Chem. 268:23964-23971(1993).
 DR EMBL; J20894; AA976834.1; -; mRNA.
 DR PIR; A49502; A49502.
 DR HSP; P18052; 1YFO.
 DR Ensembl; CG6899; Drosophila melanogaster.
 DR FlyBase; FBgn004368; Ptp4E.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0046787; F:hydrolase activity; IEA.
 DR GO; GO:0004725; F:protein tyrosine phosphatase activity; IEA.
 DR GO; GO:0004872; F:receptor activity; IEA.
 DR GO; GO:0007275; P:development; IEA.
 DR GO; GO:0006470; P:protein amino acid dephosphorylation; IEA.
 DR InterPro; IPR003961; FN III.
 DR InterPro; IPR003962; FNIII_subd.
 DR InterPro; IPR003987; TYR_phosphatase.
 DR InterPro; IPR00242; TYR_PP.
 DR Pfam; PF00041; fn3; 10.
 DR Pfam; PF00102; Y_phosphatase; 1.
 DR PRINTS; PR00014; ENTYPBIII
 DR PRINTS; PR00700; PTPPHPTASE.
 DR SMART; SM00060; FN3; 11.
 DR SMART; SM00194; PTPC; 1.
 DR PROSITE; PS00853; FN3; 11.
 DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
 DR PROSITE; PS00056; TYR_PHOSPHATASE_2; 1.
 DR PROSITE; PS00055; TYR_PHOSPHATASE_PTP; 1.
 KW Receptor.
 SQ SEQUENCE 1767 AA; 200121 MW; 70B06692B8D4B89C CRC64;

Query Match 100.0%; Score 37; DB 2; Length 1767;
 Best Local Similarity 100.0%; Pred. No. 1.2e+03; Indels 0; Gaps 0;
 Matches 7; Conservative 0; Mismatches 0;

Qy 1 ILAPVP 7
 Db 1101 ILAPVP 1107

RESULT 6

ID Q3W4F5 DROME PRELIMINARY; PRT; 1767 AA.
 AC Q9W4F5
 DT 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
 DE CG6899-PA, isoform A.
 GN Name=Ptp4E; ORFNames=CG6899;
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bertram B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Burtis K.C., Buesam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Kocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foaier C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jaitell M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.B., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milahina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacle J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=22426065; PubMed=12537568;
 RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
 RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
 RA George R.A., Hoskins R.A., Laverly T., Muzny D.M., Nelson C.R.,
 RA Pacle J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
 RA Svirskaas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
 RA Weinstein G., Scher S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
 RT "Finishing a whole-genome shotgun: release 3 of the Drosophila
 melanogaster euchromatic genome sequence.";
 RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
 RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=22426070; PubMed=12537573;
 RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J.W., Svirskaas R.,
 RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
 RA Ashburner M., Celniker S.E.;
 RT "The transposable elements of the Drosophila melanogaster euchromatin:
 RT a genomics perspective.";
 RL Genome Biol. 3:RESEARCH0084.1-RESEARCH0084.20(2002).
 RN [4]
 RP NUCLEOTIDE SEQUENCE.

RX MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Muncall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
RT systematic review.";
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
RN [5]
RP NUCLEOTIDE SEQUENCE.
RG Berkeley Drosophila Genome Project;
RA Celniker S., Carlson J., Wan K., Pfeiffer B., Frise E., George R.,
RA Hoskins R., Stapleton M., Pacleb J., Park S., Svirskas R., Smith E.,
RA Yu C., Rubin G.;
RT "Drosophila melanogaster release 4 sequence.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [6]
RP NUCLEOTIDE SEQUENCE.
RG FlyBase;
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
RN [7]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=94043220; PubMed=8226938;
RA Oon S.H., Hong A., Yang X., Chia W.;
RT "Alternative splicing in a novel tyrosine phosphatase gene (DPTP4E) of
RT Drosophila melanogaster generates two large receptor-like proteins
RT which differ in their carboxyl termini.";
RL J. Biol. Chem. 268:23964-23971(1993).
CC -|- INTERACTION;
CC Q9V6T4:CG1340; NbExp=1; IntAct=EBI-124740, EBI-98288;
CC Q9VOR3:CG17593; NbExp=1; IntAct=EBI-124740, EBI-99951;
CC Q9VX79:CG33067; NbExp=1; IntAct=EBI-124740, EBI-99044;
CC Q9VED1:CG6912; NbExp=1; IntAct=EBI-124740, EBI-102311;
CC -|- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
DR EMBL; AE003432; AAF45998.1; -; Genomic_DNA.
DR PIR; A49502; A49502.
DR PIR; B49502; B49502.
DR HSSP; P10586; 1LAR.
DR Interact; Q9W4F5; -.
DR Ensembl; CG6899; Drosophila melanogaster.
DR FlyBase; FBgn0004368; CG6899.
DR FlyBase; FBgn0004368; Ptp4E.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0004725; F:protein tyrosine phosphatase activity; IEA.
DR GO; GO:0006470; P:protein amino acid dephosphorylation; IEA.
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR003962; FNIII_subd.
DR InterPro; IPR000387; TYR_phosphatase.
DR InterPro; IPR000242; Tyr_PP.
DR Pfam; PF00041; fn3; 10_PP.
DR Pfam; PF00102; Y_phosphatase; 1.
DR PRINTS; PR00014; FNTYPEIII.
DR PRINTS; PR00700; PRTYPHPHTASE.
DR SMART; SM00060; FN3; 11.
DR SMART; SM00194; PTPC; 1.
DR PROSITE; PS50853; FN3; 11.
DR PROSITE; PS50383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50055; TYR_PHOSPHATASE_PTP; 1.
KW Hydrolase; Repeat; Transmembrane.
SQ SEQUENCE 1767 AA; 200075 MW; 4FF5CB358DB37B6C CRC64;

Query Match 100.0%; Score 37; DB 2; Length 1767;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILAPPVP 7

Db 1101 ILAPPVP 1107

RESULT 7
Q39063 ARATH PRELIMINARY; PRT; 161 AA.
ID Q39063 ARATH PRELIMINARY; PRT; 161 AA.
AC Q39063;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Cab3 promoter-binding protein.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Ecotype RLD;
RX MEDLINE=95241631; PubMed=7724680; DOI=10.1104/pp.107.2.657;
RA Mitra A., Higgins D.W.;
RT "Nucleotide sequence of a cDNA clone from Arabidopsis thaliana
RT encoding a small A-T tract-binding protein.";
RL Plant Physiol. 107:657-658(1995).
DR EMBL; L33781; AAA87293.1; -; mRNA.
SQ SEQUENCE 161 AA; 18479 MW; 0C1D81B7C9226B73 CRC64;

Query Match 97.3%; Score 36; DB 2; Length 161;
Best Local Similarity 85.7%; Pred. No. 1.5e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILAPPVP 7
Db 123 VLAPPVP 129

RESULT 8
Q6ZN89 HUMAN PRELIMINARY; PRT; 218 AA.
ID Q6ZN89 HUMAN PRELIMINARY; PRT; 218 AA.
AC Q6ZN89;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein FLJ16328.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Synovial membrane tissue;
RA Suzuki O., Sasaki N., Aotsuka S., Shoji T., Ichihara T., Shiohata N.,
RA Matsumoto K., Hirano M., Sano S., Nomura R., Yoshikawa Y.,
RA Matsumura Y., Moriya S., Chiba E., Momiyama H., Onogawa S.,
RA Kaeriyama S., Satoh N., Matsunawa H., Takahashi E., Kataoka R.,
RA Kuga N., Kuroda A., Satoh I., Kamata K., Takami S., Terashima Y.,
RA Watanabe M., Sugiyama T., Irie R., Otsuki T., Sato H., Wakamatsu A.,
RA Ishii S., Yamamoto J., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T.,
RA Kimura K., Yamashita H., Matsuo K., Nakamura Y., Sekine M.,
RA Kikuchi H., Kanda K., Wagatsuma M., Murakawa K., Kanehori K.,
RA Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B., Suzuki Y.,
RA Sugano S., Nagahari K., Masuho Y., Nagai K., Isogai T.;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK131326; BADI8486.1; -; mRNA.
SQ SEQUENCE 218 AA; 23633 MW; 12D112E51063B4A8 CRC64;

Query Match 97.3%; Score 36; DB 2; Length 218;
Best Local Similarity 85.7%; Pred. No. 2e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILAPPVP 7

Db 93 VLAPPVP 99

```

RESULT 9
Q7SBK3_NEUCR
ID Q7SBK3_NEUCR PRELIMINARY; PRT; 461 AA.
AC Q7SBK3;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Predicted protein.
GN Name=NCU08547.1;
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=OR74A;
RA Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,
RA Jaffe D., FitzHugh W., Ma L.-J., Smirnov S., Purcell S., Rehman B.,
RA Elkins T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,
RA Qui D., Ianakiev P., Pedersen D., Nelson M., Washburne M.,
RA Selitrennikoff C.P., Kinsey J.A., Braun E.L., Zelter A., Schulte U.,
RA Kothe G.O., Jedd G., Mewes W., Staben C., Marcotte E., Greenberg D.,
RA Roy A., Foley K., Naylor J., Thomann N., Barrett R., Gnerre S.,
RA Kamal M., Kamysellis M., Mauceli E., Bielke C., Rudd S., Frishman D.,
RA Kryatova S., Rasmussen C., Metzberg R.L., Perkins D.D., Kroken S.,
RA Cogoni C., Macino G., Catcheside D., Li W., Pratt R.J., Osmani S.A.,
RA DeSouza C.C., Glass L., Orbach M.J., Berglund J., Voelker R.,
RA Yarden O., Plamann M., Seiler S., Dunlap J., Radford A., Aramayo R.,
RA Natvig D.O., Alex L.A., Mannhaupt G., Ebbole D.J., Freitag M.,
RA Paulsen I., Sachs M.S., Lander E.S., Nussbaum C., Birren B.;
RA "The Genome Sequence of the Filamentous Fungus Neurospora crassa."
RL Nature 0:0-0(2003).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC EMBL; AABX01000166; EAA33793.1; -; Genomic DNA.
SQ SEQUENCE 461 AA; 51895 MW; F6965A30AAF31089 CRC64;

Query Match 97.3%; Score 36; DB 2; Length 461;
Best Local Similarity 85.7%; Pred. No. 4.5e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILAPPVP 7
Db 424 VLAPPVP 430

RESULT 10
Q6Z8U5_ORYSA
ID Q6Z8U5_ORYSA PRELIMINARY; PRT; 75 AA.
AC Q6Z8U5;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Hypothetical protein P0686H1.1 (Hypothetical protein
DE P0605H02.41).
GN Name=P0686H1.1; Synonyms=P0605H02.41;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzoideae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsunoto T., Yamamoto K.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsunoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 8, PAC
RT clone:P0605H02.",

Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP004762; BAD09996.1; -; Genomic DNA.
DR EMBL; AP004620; BAD09674.1; -; Genomic DNA.
DR Gramene; Q6Z8U5; -.
KW Hypothetical protein.
SQ SEQUENCE 75 AA; 8928 MW; 0542C39BEC9DF68 CRC64;

Query Match 94.6%; Score 35; DB 2; Length 75;
Best Local Similarity 71.4%; Pred. No. 98;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILAPPVP 7
Db 64 VLAPPVP 70

RESULT 11
Q5D5E5_9RICK PRELIMINARY; PRT; 186 AA.
AC Q5D5E5;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Phenylacrylic acid decarboxylase, 3-octaprenyl-4-hydroxybenzoate
DE carboxy-lyase (EC 4.1.1.1.-).
GN Name=ubix; ORFNames=WMoj0053;
OS Wolbachia endosymbiont of Drosophila mojavensis.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Rickettsiaceae; Wolbachieae; Wolbachia.
OX NCBI_TaxID=307503;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Salzberg S.L., Dunning Hotopp J.C., Delcher A.L., Pop M., Smith D.R.,
RA Eisen M.B., Nelson W.C.;
RT "Serendipitous discovery of Wolbachia genomes in multiple Drosophila
RT species".
RL Genome Biol. 6:R23-R23(2005).
DR EMBL; AY897517; AAX14444.1; -; Genomic DNA.
DR InterPro; IPR003382; Flavoprotein.
DR InterPro; IPR004507; PAD1.
DR Pfam; PF02441; Flavoprotein; 1.
DR TIGRPFAM; TIGR00421; ubix_pad; 1.
KW Lyase.
SQ SEQUENCE 186 AA; 20551 MW; 95B08191CBD0141E CRC64;

Query Match 94.6%; Score 35; DB 2; Length 186;
Best Local Similarity 85.7%; Pred. No. 2.6e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILAPPVP 7
Db 144 ILAPPVP 150

RESULT 12
Q4UL63_RICFE PRELIMINARY; PRT; 191 AA.
AC Q4UL63;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE 3-octaprenyl-4-hydroxybenzoate carboxy-lyase (EC 4.1.1.1.-).
GN Name=ubix; OrderedLocNames=RP_0859;
OS Rickettsia felis (Rickettsia azadi).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Rickettsiaceae; Rickettsieae; Rickettsia; spotted fever group.
OX NCBI_TaxID=42862;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=URRWXCal2;
RX PubMed=15984913; DOI=10.1371/journal.pbio.0030248;
RA Ogata H., Renesto P., Audic S., Robert C., Blanc G., Fournier P.-E.,
RA Parinello H., Claverie J.-M., Raoult D.;

```



```

RT "The genome sequence of Rickettsia felis identifies the first putative
RT conjugative plasmid in an obligate intracellular parasite.";
RL PLoS Biol. 3:E248-E248(2005).
DR EMBL; CP000053; AAY61710.1; -; Genomic_DNA.
DR InterPro; IPR003382; Flavoprotein.
DR InterPro; IPR004507; PAD1.
DR Pfam; PF02441; Flavoprotein; 1.
DR TIGRfams; TIGR00421; ubix_pad; 1.
KW Complete proteome; Lyase.
SQ SEQUENCE 191 AA; 21189 MW; 6D1C054513DE8582 CRC64;

Query Match 94.6%; Score 35; DB 2; Length 191;
Best Local Similarity 85.7%; Pred. No. 2.6e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAPVP 7
Db 146 ILAPVP 152

RESULT 13
Q73HK2 WOLPM
ID Q73HK2 WOLPM PRELIMINARY; PRT; 191 AA.
AC Q73HK2;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Phenylacrylic acid decarboxylase, 3-octaprenyl-4-hydroxybenzoate
DE carboxy-lyase (EC 4.1.1.-).
GN Name=ubix; OrderedLocNames=WD0556;
OS Wolbachia pipitensis wMel.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Rickettsiaceae; Wolbachiae; Wolbachia.
OX NCBI_TaxID=66077;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX PubMed=15024419; DOI=10.1371/journal.pbio.0020069;
RA Wu M., Sun L.V., Vamathevan J.J., Kiegler M., DeBoy R.T.,
RA Brownlie J.C., McGraw E.A., Martin W., Esser C., Almadinejad N.,
RA Wiegand C., Madupu R., Beanan M.J., Brinkac L.M., Daugherty S.C.,
RA Durkin A.S., Kolonay J.F., Nelson M.W.C., Mohamoud Y., Lee P.,
RA Berry K.J., Young M.B., Utterback T.R., Weidman J.F., Nierman W.C.,
RA Paulsen I.T., Nelson K.E., Tettelin H., O'Neill S.L., Eisen J.A.;
RA "Phylogenomics of the reproductive parasite Wolbachia pipitensis wMel:
RT a streamlined genome overrun by mobile genetic elements.";
RL PLoS Biol. 2:327-341(2004).
DR EMBL; AE017258; AAS14263.1; -; Genomic_DNA.
DR TIGR; WD0556; -; F:carboxy-lyase activity; IEA.
DR GO; GO:0016831; F:carboxy-lyase activity; IEA.
DR InterPro; IPR003382; Flavoprotein.
DR InterPro; IPR004507; PAD1.
DR Pfam; PF02441; Flavoprotein; 1.
DR TIGRfams; TIGR00421; ubix_pad; 1.
KW Complete proteome; Lyase.
SQ SEQUENCE 191 AA; 21146 MW; FE4BE4EA173556DB CRC64;

Query Match 94.6%; Score 35; DB 2; Length 191;
Best Local Similarity 85.7%; Pred. No. 2.6e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAPVP 7
Db 144 ILAPVP 150

RESULT 14
Q5PAP8 ANAMM
ID Q5PAP8 ANAMM PRELIMINARY; PRT; 201 AA.
AC Q5PAP8;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE 3-octaprenyl-4-hydroxybenzoate carboxyl-lyase (EC 4.1.1.-).

```

```

GN Name=ubix; OrderedLocNames=AM645;
OS Anaplasma marginale (strain St. Maries).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Anaplasmataceae; Anaplasma.
OX NCBI_TaxID=234826;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX PubMed=15618402; DOI=10.1073/pnas.0406656102;
RA Brayton K.A., Kappmeyer L.S., Herndon D.R., Dark M.J., Tibbals D.L.,
RA Palmer G.H., McGuire T.C., Knowles D.P. Jr.;
RT "Complete genome sequencing of Anaplasma marginale reveals that the
RT surface is skewed to two superfamilies of outer membrane proteins.";
RL Proc. Natl. Acad. Sci. U.S.A. 102:844-849(2005).
DR EMBL; CP000030; AAV86632.1; -; Genomic_DNA.
DR GO; GO:0016831; F:carboxy-lyase activity; IEA.
DR GO; GO:0016829; F:lyase activity; IEA.
DR InterPro; IPR003382; Flavoprotein.
DR InterPro; IPR004507; PAD1.
DR Pfam; PF02441; Flavoprotein; 1.
DR TIGRfams; TIGR00421; ubix_pad; 1.
KW Complete proteome; Lyase.
SQ SEQUENCE 201 AA; 22035 MW; 3FAFFA922641699A CRC64;

Query Match 94.6%; Score 35; DB 2; Length 201;
Best Local Similarity 85.7%; Pred. No. 2.8e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAPVP 7
Db 159 ILAPVP 165

RESULT 15
Q60W03 CAEBR
ID Q60W03 CAEBR PRELIMINARY; PRT; 231 AA.
AC Q60W03;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein CBG19323 (Fragment).
GN Name=CBG19323;
OS Caenorhabditis briggsae.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6238;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RG The C.briggsae Sequencing Consortium;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CAAC01000098; CAE72212.1; -; Genomic_DNA.
DR GO; GO:0005199; F:structural constituent of cell wall; IEA.
DR InterPro; IPR003882; PstII_extensin.
DR PRINTS; PR01218; PSTIEXTENSIN.
KW Hypothetical protein.
FT NON TER 1
SQ SEQUENCE 231 AA; 21940 MW; DC7DD1E86F46C959 CRC64;

Query Match 94.6%; Score 35; DB 2; Length 231;
Best Local Similarity 85.7%; Pred. No. 3.2e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAPVP 7
Db 112 ILAPVP 118

RESULT 16
Q7D5X3 MYCTU
ID Q7D5X3 MYCTU PRELIMINARY; PRT; 266 AA.
AC Q7D5X3;

```

DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DE 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE Soj family protein.
GN OrderedLocusNames=MT3309;
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium tuberculosis complex.
OX NCBI_TaxID=1773;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CDC 1551 / Oshkosh;
RX MEDLINE=22206494; PubMed=12218036;
DOI=10.1128/JB.184.19.5479-5490.2002;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J.D., DeBoy R.T., Dodson R.J., Gwinn M.L., Haft D.H.,
RA Hickey E.K., Kolonay J.P., Nelson W.C., Umayam L.A., Ermolaeva M.D.,
RA Salzberg S.L., Delcher A., Uterback T.R., Weidman J.F., Khouri H.M.,
RA Gill J., Mikula A., Bishai W., Jacobs W.R. Jr., Venter J.C.,
RA Fraser C.M.;
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains";
RL J. Bacteriol. 184:5479-5490(2002).
DR EMBL; AE000516; AAK47651.1; -; Genomic_DNA.
DR TIGR; MT3309; -;
DR GO; GO:0042242; F:cobyrinic acid a,c-diamide synthase activity; IEA.
DR GO; GO:0009236; P:cobalamin biosynthesis; IEA.
DR InterPro; IPR002586; CbiA_P_synth.
DR Pfam; PF01656; CbiA; 1.
SQ SEQUENCE 266 AA; 28576 MW; 49BACF4E0ECC2116 CRC64;

Query Match 94.6%; Score 35; DB 2; Length 266;
Best Local Similarity 71.4%; Pred. No. 3.8e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAPPVP 7
:|||||:
DB 209 VLAPPPI 215

RESULT 17
ID C05853 MYCTU PRELIMINARY; PRT; 266 AA.
AC C05853;
DT 01-JUL-1997 (TReMBLrel. 04, Created)
DT 01-JUL-1997 (TReMBLrel. 04, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE POSSIBLE SOJ/PARA-RELATED PROTEIN.
GN OrderedLocusNames=RV3213C;
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium tuberculosis complex.
OX NCBI_TaxID=1773;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230; DOI=10.1038/31159;
RA Cole S.T., Broach R., Parkhill J., Garnier T., Churcher C.M.,
RA Harris D.E., Gordon S.V., Eigmeier K., Gas S., Barry C.E. III,
RA Tekala F., Badcock K., Basham D., Brown D., Chillingworth T.,
RA Connor R., Davies R.M., Devlin K., Feltwell T., Gentles S., Hamlin N.,
RA Holroyd S., Hornsby T., Jagels K., Krogh A., McLean J., Moule S.,
RA Murphy L.D., Oliver S., Osborne J., Quail M.A., Rajandream M.A.,
RA Rogers J., Rutter S., Seeger K., Skelton S., Squares R.,
RA Sultston J.B., Taylor K., Whitehead S., Barrall B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence";
RL Nature 393:537-544(1998).
DR EMBL; BX842582; CAB08303.1; -; Genomic_DNA.
DR PIR; P70595; P70595.
DR TubercuList; RV3213C; -.

DR GO; GO:0042242; F:cobyrinic acid a,c-diamide synthase activity; IEA.
DR GO; GO:0009236; P:cobalamin biosynthesis; IEA.
DR InterPro; IPR002586; CbiA_P_synth.
DR Pfam; PF01656; CbiA; 1.
KW Complete proteome.
SQ SEQUENCE 266 AA; 28575 MW; 419AC96E00CC2116 CRC64;

Query Match 94.6%; Score 35; DB 2; Length 266;
Best Local Similarity 71.4%; Pred. No. 3.8e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAPPVP 7
:|||||:
DB 209 VLAPPPI 215

RESULT 18
ID Q73UQ2 MYCPA PRELIMINARY; PRT; 266 AA.
AC Q73UQ2;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=MAP3314C;
OS Mycobacterium paratuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium avium complex (MAC).
OX NCBI_TaxID=1770;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=k10;
RA Li L., Bannantine J., Zhang Q., Amonsin A., Alt D., Kapur V.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE017239; AAS05864.1; -; Genomic DNA.
DR GO; GO:0042242; F:cobyrinic acid a,c-diamide synthase activity; IEA.
DR GO; GO:0009236; P:cobalamin biosynthesis; IEA.
DR InterPro; IPR002586; CbiA_P_synth.
DR Pfam; PF01656; CbiA; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 266 AA; 28104 MW; 429E782A29C7781A CRC64;

Query Match 94.6%; Score 35; DB 2; Length 266;
Best Local Similarity 71.4%; Pred. No. 3.8e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAPPVP 7
:|||||:
DB 209 VLAPPPI 215

RESULT 19
ID Q7TWZ9 MYCBO PRELIMINARY; PRT; 266 AA.
AC Q7TWZ9;
DT 01-OCT-2003 (TReMBLrel. 25, Created)
DT 01-OCT-2003 (TReMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE POSSIBLE SOJ/PARA-RELATED PROTEIN.
GN OrderedLocusNames=Mb3239C;
OS Mycobacterium bovis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium tuberculosis complex.
OX NCBI_TaxID=1765;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=AF2122/97;
RX MEDLINE=22709107; PubMed=12788972; DOI=10.1073/pnas.1130426100;
RA Garnier T., Eigmeier K., Camus J.-C., Medina N., Mansoor H.,
RA Pryor M., Dutchoy S., Grondin S., Lacroix C., Monsemp C., Simon S.,
RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,

```
RA Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RT "The complete genome sequence of Mycobacterium bovis.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
DR EMBL; BX248345; CAD95331.1; -; Genomic DNA.
DR GO; GO:0042242; F:coybinic acid a,c-diamide synthase activity; IEA.
DR GO; GO:0009236; P:cobalamin biosynthesis; IEA.
DR InterPro; IPR002586; CbiA_P_synth.
DR Pfam; PF01656; CbiA; 1.
KW Complete proteome.
SQ SEQUENCE 266 AA; 28576 MW; 49BACFAE08CC2116 CRC64;

Query Match 94.6%; Score 35; DB 2; Length 266;
Best Local Similarity 71.4%; Pred. No. 3.8e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAPPVP 7
:|||||
Db 209 VLAPPVP 215

RESULT 20
Q5WMS2_ORYSA
ID Q5WMS2_ORYSA PRELIMINARY; PRT; 477 AA.
AC Q5WMS2;
DT 25-OCT-2004 (TReMBLrel. 28, Created)
DT 25-OCT-2004 (TReMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TReMBLrel. 28, Last annotation update)
DE Hypothetical protein OJ1333_C12.9.
GN Name=OJ1333_C12.9;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Chow T.-Y., Haing Y.-J.C., Chen C.-S., Chen H.-H., Liu S.-M.,
RA Chao Y.-T., Chang S.-J., Chen H.-C., Chen S.-K., Chen T.-R.,
RA Chen Y.-L., Cheng C.-H., Chung C.-I., Han S.-Y., Hsiao S.-H.,
RA Hsiung J.-N., Hsu C.-H., Huang J.-J., Kau P.-I., Lee M.-C.,
RA Li Y.-P., Lin S.-J., Lin Y.-C., Wu S.-W., Yu C.-Y., Yu S.-W.,
RA Wu H.-P., Shaw J.-F., *Yu Y., *Rambo T., *Currie J., *Collura K.,
RA *Soderlund C., *Wing R.;
RT "Oryza sativa BAC OJ1333; C12 genomic sequence.";
RL Submitted (OCT-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC104712; AA032163.1; -; Genomic DNA.
DR InterPro; IPR003480; Transferase.
DR Pfam; PF02458; Transferase; 1.
KW Hypothetical protein.
SQ SEQUENCE 477 AA; 52057 MW; 44E7F43BD9637C96 CRC64;

Query Match 94.6%; Score 35; DB 2; Length 477;
Best Local Similarity 85.7%; Pred. No. 7e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAPPVP 7
:|||||
Db 127 LIAPPVP 133

RESULT 21
Q7PYUO_ANOGA
ID Q7PYUO_ANOGA PRELIMINARY; PRT; 662 AA.
AC Q7PYU0;
DT 01-MAR-2004 (TReMBLrel. 26, Created)
DT 01-MAR-2004 (TReMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE ENSANGP0000017547 (Fragment).
GN ORFNames=ENSANGG00000015058;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Culicidae;
OC Anophelinae; Anopheles.
```

```
OX NCBI_TaxID=180454;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAINS-PEST;
RA Anopheles Genome Sequencing Consortium;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAAB01008987; EAA01467.1; -; Genomic DNA.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003684; F:damaged DNA binding; IEA.
DR GO; GO:0006289; P:nucleotide-excision repair; IEA.
DR InterPro; IPR002345; Lipocalin.
DR InterPro; IPR004583; Rad4.
DR PANTHER; PTHR12135; Rad4; 1.
DR Pfam; PF03835; Rad4; 1.
DR PROSITE; PS00213; LIPOCALIN; UNKNOWN_1.
FT NON TER 1
SQ SEQUENCE 662 AA; 76575 MW; 0B18D6974A185D87 CRC64;

Query Match 94.6%; Score 35; DB 2; Length 662;
Best Local Similarity 85.7%; Pred. No. 9.9e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAPPVP 7
:|||||
Db 141 LIAPPVP 147

RESULT 22
Q55ZY4_CRYNE
ID Q55ZY4_CRYNE PRELIMINARY; PRT; 842 AA.
AC Q55ZY4;
DT 13-SEP-2005 (TReMBLrel. 31, Created)
DT 13-SEP-2005 (TReMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TReMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=CNBA3560;
OS Cryptococcus neoformans var. neoformans B-3501A.
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Heterobasidiomycetes;
OC Tremellomycetidae; Tremellales; Tremellaceae; Filobasidiella.
OX NCBI_TaxID=283643;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=B-3501A;
RA Fung E., Hyman R.W., Rowley D., Bruno D., Miranda M., Fukushima M.,
RA Wickes B.L., Fu J., Davis R.W.;
RT "Cryptococcus neoformans serotype D sequencing.";
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AA01000002; EAL23240.1; -; Genomic DNA.
KW Hypothetical protein.
SQ SEQUENCE 842 AA; 94227 MW; DA8051D58EC1ABE2 CRC64;

Query Match 94.6%; Score 35; DB 2; Length 842;
Best Local Similarity 71.4%; Pred. No. 1.3e+03;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAPPVP 7
:|||||
Db 200 VLAPPVP 206

RESULT 23
Q5KP92_CRYNE
ID Q5KP92_CRYNE PRELIMINARY; PRT; 842 AA.
AC Q5KP92;
DT 10-MAY-2005 (TReMBLrel. 30, Created)
DT 10-MAY-2005 (TReMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TReMBLrel. 30, Last annotation update)
```

OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
 OX NCBI_TaxID=44270;
 [1]
 RN NUCLEOTIDE SEQUENCE.
 RC STRAIN=MNYC/BZ/62/M379;
 RA Gokool S.;
 RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
 RR EMBL; AY170466; AA017687.1; -; Genomic_DNA.
 DR InterPro; IPR002553; Adaptin_N.
 DR InterPro; IPR002088; PPTA.
 DR InterPro; IPR008162; Pyrophosphatase.
 DR Pfam; PF01602; Adaptin_N; 1.
 DR PROSITE; PS00387; PPTASE; UNKNOWN 1.
 DR PROSITE; PS00904; PPTA; UNKNOWN 1.
 SQ SEQUENCE 990 AA; 108292 MW; 346B7BDB80F9A013 CRC64;
 Query Match 94.6%; Score 35; DB 2; Length 990;
 Best Local Similarity 71.4%; Pred. No. 1.5e+03;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ILAPPVP 7
 :|||||
 Db 859 VLAPPIP 865
 RESULT 25
 Q4QGY0 LEIMA
 ID Q4QGY0 LEIMA PRELIMINARY; PRT; 990 AA.
 AC Q4QGY0;
 DT 13-SEP-2005 (TrEMBLrel. 31, Created)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
 DE Adaptin-related protein-like protein.
 GN OSFNames=LmjF11.0990;
 GN Leishmania major.
 OS Leishmania major.
 OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
 OX NCBI_TaxID=5664;
 [1]
 RN NUCLEOTIDE SEQUENCE.
 RC STRAIN=Friedlin;
 RA Peacock C.S.; Murphy L.; Ivens A.C.; Berriman M.; Blackwell J.;
 RA Smith D.; Collins M.; Foster N.; Harris D.; Oliver K.; O'Neill S.;
 RA Saunders D.; Seeger K.; Warren T.; Rajandream M.; and Barrell B.G.;
 RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
 DR EMBL; CTO05250; CAJ02880.1; -; Genomic DNA.
 SQ SEQUENCE 990 AA; 108501 MW; 8D752523AEA39F7F CRC64;
 Query Match 94.6%; Score 35; DB 2; Length 990;
 Best Local Similarity 71.4%; Pred. No. 1.5e+03;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ILAPPVP 7
 :|||||
 Db 860 VLAPPIP 866
 RESULT 26
 PAAD RICPR
 ID PAAD RICPR STANDARD; PRT; 189 AA.
 AC Q9ZD09;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 10-MAY-2005 (Rel. 47, Last annotation update)
 DE Probable aromatic acid decarboxylase [EC 4.1.1.-].
 GN OrderedLocusNames=RP541;
 GN Rickettsia prowazekii.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
 OC Rickettsiaceae; Rickettsia; typhus group.
 OX NCBI_TaxID=782;
 [1]
 RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=Madrid E;
 RX MEDLINE=99039499; PubMed=9823893; DOI=10.1038/24094;

```
RA Andersson S.G.E., Zomorodipour A., Andersson J.O.,
RA Sicheritz-Ponten T., Alemark U.C.M., Podowski R.M., Naeslund A.K.,
RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
RT "The genome sequence of Rickettsia prowazekii and the origin of
RT mitochondria.";
RL Nature 396:133-140(1998).
CC -!- SIMILARITY: Belongs to the polyprenyl P-hydroxybenzoate /
CC phenylacrylic acid decarboxylases family.
CC -----
CC This Swiss-Prot entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; AJ235272; CAA14990.1; -; Genomic_DNA.
DR PIR; D71658; D71658.
DR InterPro; IPR003382; Flavoprotein.
DR InterPro; IPR004507; PAD1.
DR Pfam; PF02441; Flavoprotein; 1.
DR TIGRFAMs; TIGR00421; ubix_pad; 1.
KW Complete proteome; Decarboxylase; Hypothetical protein; Lyase.
SQ SEQUENCE 189 AA; 20973 MW; DD2DBF374C30B8E6 CRC64;

Query Match 91.9%; Score 34; DB 1; Length 189;
Best Local Similarity 71.4%; Pred. No. 3.9e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAPPVP 7
DB 146 ILAPPVP 152

"
RESULT 27
Q68WJ3 RICTY
ID Q68WJ3 RICTY PRELIMINARY; PRT; 190 AA.
AC Q68WJ3.
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE 3-octaprenyl-4-hydroxybenzoate carboxy-lyase (EC 4.1.1.-).
GN Name=ubix; OrderedLocustNames=RT0530;
OS Rickettsia typhi.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Rickettsiaceae; Rickettsiae; Rickettsia; typhus group.
OX NCBI_TaxID=785;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX STRAIN=Wilmington;
RX PubMed=15317790; DOI=10.1128/JB.186.17.5842-5855.2004;
RA McLeod M.P., Qin X., Karpathy S.E., Gioia J., Highlander S.K.,
RA Fox G.E., McNeill T.Z., Jiang H., Muzny D., Jacob L.S., Hawes A.C.,
RA Sodergren E., Gill R., Hume J., Morgan M., Fan G., Amin A.G.,
RA Gibbs R.A., Hong C., Yu X.-J., Walker D.H., Weinstein G.M.;
RT "Complete genome sequence of Rickettsia typhi and comparison with
RT sequences of other Rickettsiae.";
RL J. Bacteriol. 186:5842-5855(2004).
DR EMBL; AE017197; RAU03999.1; -; Genomic DNA.
DR GO; GO:0016829; F:lyase activity; IEA.
DR InterPro; IPR003382; Flavoprotein.
DR InterPro; IPR004507; PAD1.
DR Pfam; PF02441; Flavoprotein; 1.
DR TIGRFAMs; TIGR00421; ubix_pad; 1.
KW Complete proteome; Lyase.
SQ SEQUENCE 190 AA; 20858 MW; B2E82B16BDD2A67 CRC64;

Query Match 91.9%; Score 34; DB 2; Length 190;
Best Local Similarity 71.4%; Pred. No. 4e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAPPVP 7
DB 146 ILAPPVP 152
```

```
RESULT 28
Q62I56 BURMA
ID Q62I56 BURMA PRELIMINARY; PRT; 315 AA.
AC Q62I56;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Transcriptional regulator, LysR family.
OS OrderedLocustNames=BMA2032;
GN Burkholderia mallei (Pseudomonas mallei).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Burkholderia.
OX NCBI_TaxID=13373;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX STRAIN=ATCC 23344;
RX PubMed=15377793; DOI=10.1073/pnas.0403306101;
RA Nierman W.C., DeShazer D., Kim H.S., Tettelin H., Nelson K.E.,
RA Feldblyum T.V., Ulrich R.L., Ronning C.M., Brinkac L.M.,
RA Daugherty S.C., Davidse T.D., Deboy R.T., Dimitrov G., Dodson R.J.,
RA Durkin A.S., Gwinn M.L., Haft D.H., Khouri H.M., Kolonay J.F.,
RA Madupu R., Mohammed Y., Nelson W.C., Radune D., Romero C.M.,
RA Sarria S., Selengut J., Shamblin C., Sullivan S.A., White O., Yu Y.,
RA Zafar N., Zhou L., Fraser C.M.;
RT "Structural flexibility in the Burkholderia mallei genome.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:14246-14251(2004).
CC -!- SIMILARITY: Contains 1 HTH LysR-type DNA-binding domain.
DR EMBL; CP000010; AAU49569.1; -; Genomic_DNA.
DR TIGR; BMA2032; -;
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR InterPro; IPR000847; HTH_LysR.
DR InterPro; IPR005119; LysR_subst_bd.
DR InterPro; IPR011991; Wing_hlx_DNA_bd.
DR Pfam; PF00126; HTH_1; 1.
DR Pfam; PF03466; LysR_substrate; 1.
DR PRINTS; PR00039; HTHLYSR.
DR PROSITE; PS0931; HTH_LYSR; 1.
KW Activator; Complete proteome; DNA-binding; Transcription;
KW Transcription regulation.
SQ SEQUENCE 315 AA; 34879 MW; 8BAD4282598CD429 CRC64;

Query Match 91.9%; Score 34; DB 2; Length 315;
Best Local Similarity 71.4%; Pred. No. 6.8e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAPPVP 7
DB 268 VIAPPVP 274

"
RESULT 29
Q5RI92 BRARE
ID Q5RI92 BRARE PRELIMINARY; PRT; 341 AA.
AC Q5RI92;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Novel protein similar to prostaglandin receptor E2 (Subtype EP2)
DE (PTGER2).
GN Name=OTTDARP00000003831; ORFNames=CH211-13N20.1-001;
GN Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Hunter G.;
RL Submitted (DEC-2004) to the EMBL/GenBank/DBJ databases.
```

-1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
EMBL: BX284679; CAIL1598.1; -, Genomic DNA.
DR GO: GO:0006021; C: integral to membrane; IEA.
DR GO: GO:0004872; F: receptor activity; IEA.
DR GO: GO:0001584; F: rhodopsin-like receptor activity; IEA.
DR GO: GO:0004960; F: thromboxane receptor activity; IEA.
DR GO: GO:0007186; P: G-protein coupled receptor protein signalin. . .; IEA.
DR GO: GO:0007165; P: signal transduction; IEA.
DR InterPro: IPR002776; GPCR_rhodopsin.
DR InterPro: IPR000370; IP_receptor.
DR InterPro: IPR008365; Prostanoid_recept.
DR InterPro: IPR001105; Tbox_receptor.
DR Pfam: PF00001; 7tm 1; 1.
DR PRINTS: PR00237; GPCRHOOPS.
DR PRINTS: PR01788; PROSTANOIDR.
DR PRINTS: PR00856; PRSTNOIDIPR.
DR PRINTS: PR00429; THROMBOXANER.
DR PROSITE: PS00237; G_PROTEIN_RECP_F1_1; UNKNOWN_1.
DR PROSITE: PS0262; G_PROTEIN_RECP_F1_2; 1.
KW G-protein coupled receptor; Receptor; Transducer; Transmembrane.
SQ SEQUENCE 341 AA; 38790 MW; C8F1C4D5DA6981CD CRC64;

Query Match 91.9%; Score 34; DB 2; Length 341;
Best Local Similarity 85.7%; Pred. No. 7.4e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILAPPVP 7

Db 300 ILSPVP 306

RESULT 30
Q4SLJ8 TETNG
ID Q4SLJ8 TETNG PRELIMINARY; PRT; 467 AA.
AC Q4SLJ8
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Chromosome 15 SCAP1556, whole genome shotgun sequence.
GN ORFNames=GSTENG00016208001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
RN NUCLEOTIDE SEQUENCE.
RP Jaillon O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Dasilva C., Salancoubat M., Levy M., Boudet N., Castellano S.,
RA Bionhoud V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA Biemont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
RA Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Guzy J.,
RA Parra G., Lardier G., Chapelle C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Volff J.N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Winkler P., Lander E.S., Weissbach J., Roest Crollius H.,
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
the early vertebrate proto-karyotype."
RL Nature 431:946-957(2004).
RN NUCLEOTIDE SEQUENCE.
RP Genoscope; Whitehead Institute Centre for Genome Research;
RG Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
preliminary data.
DR EMBL; CAEE01014556; CAF98484.1; -, Genomic DNA.
SQ SEQUENCE 467 AA; 49745 MW; 9517E9DB4F0E6D37 CRC64;

Query Match 91.9%; Score 34; DB 2; Length 467;
Best Local Similarity 71.4%; Pred. No. 1e+03;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILAPPVP 7

Db 352 LLAPPVP 358

RESULT 31
Q4IM73 GIBZE
ID Q4IM73 GIBZE PRELIMINARY; PRT; 621 AA.
AC Q4IM73
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=FG01685.1;
OS Gibberella zeae PH-1.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
OX NCBI_TaxID=229533;
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=PH-1;
RA Birren B., Nusbaum C., Abouelleil A., Allen N., Anderson S.,
RA Arachchi H.M., Barna N., Bastien V., Bloom T., Boguslavskiy L.,
RA Boukhgalter B., Butler J., Calvo S.E., Camarata J., Chang J.,
RA Choepel Y., Collymore A., Cook A., Cooke P., Corum B., DeArelano K.,
RA Diaz J.S., Dodge S., Dooley K., Dorril L., Elkins T., Engels R.,
RA Erickson J., Faro S., Ferreira P., Fitzgerald M., Gage D., Galagan J.,
RA Gardyna S., Gherre S., Graham L., Grand-Pierre N., Hafez N.,
RA Hagopian D., Hagos B., Hall J., Horton L., Hulme W., Iliev I.,
RA Jaffe D., Johnson R., Jones C., Kamal M., Kamat A., Karatas A.,
RA Kells C., Landers T., Levine R., Lindblad-Toh K., Liu G., Lui A.,
RA Ma L.-J., Mabbitt R., MacLean C., Macdonald P., Major J., Manning J.,
RA Matthews C., Mauceli E., McCarthy M., Meldrim J., Meneus L.,
RA Mihova T., Mienga V., Murphy T., Naylor J., Nguyen C., Nicol R.,
RA Nielsen C.B., Norbu C., O'Connor T., O'Donnell P., O'Neill D.,
RA Oliver J.B., Peterson K., Phunkhang P., Pierre N., Purcell S.,
RA Rachupka A., Ramasamy U., Raymond C., Rietz R., Rise C., Rogov P.,
RA Roman J., Schauer S., Schupbach R., Seaman S., Severy P., Smirnov S.,
RA Smith C., Spencer B., Stange-Thomann N., Stojanovic N., Stubbs M.,
RA Talamas J., Tesfaye S., Theodore J., Topham K., Travers M.,
RA Vasiliev H., Venkataraman V.S., Viel R., Vo A., Wang S., Wilson B.,
RA Wu X., Wyman D., Young G., Zainoun J., Zembek L., Zimmer A., Zody M.,
RA Lander E.;
RT "Fusarium graminearum genome sequence."
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
preliminary data.
DR EMBL; AACM0100090; EAA67830.1; -, Genomic DNA.
KW Hypothetical protein.
SQ SEQUENCE 621 AA; 70293 MW; 7CC8B3E25713A93B CRC64;

Query Match 91.9%; Score 34; DB 2; Length 621;
Best Local Similarity 71.4%; Pred. No. 1.4e+03;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILAPPVP 7

Db 451 VIAPPVP 457

RESULT 32
Q8T3T1 LYTVTA
ID Q8T3T1 LYTVTA PRELIMINARY; PRT; 762 AA.
AC Q8T3T1
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Major yolk protein (Fragment).

OS *Lytechinus variegatus* (Sea urchin).
 OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
 OC Echinoidea; Euechinoidea; Echinacea; Temnopleuroidea; Toxopneustidae;
 OC Lytechinus.
 OX NCBI_TaxID=7654;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=21966081; PubMed=11969251; DOI=10.1006/dbio.2002.0611;
 RA Brooks J.M., Wessel G.M.;
 RT "The major yolk protein in sea urchins is a transferrin-like, iron
 binding protein.";
 RL Dev. Biol. 245:1-12(2002).
 DR EMBL; AY090113; AAM14718.1; -; mRNA.
 DR HSSP; P56410; 10VW.
 DR GO; GO:0005576; C:extracellular region; IEA.
 DR GO; GO:0008199; F:ferric iron binding; IEA.
 DR GO; GO:0006879; P:iron ion homeostasis; IEA.
 DR GO; GO:0006826; P:iron ion transport; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR001156; Peptidase_S60.
 DR InterPro; IPR008197; WAP.
 DR Pfam; PF00405; Transferrin; 1.
 DR Pfam; PF00095; WAP; 1.
 DR PRINTS; PR00422; TRANSFERRIN.
 DR SMART; SM00094; TR_FER; 1.
 DR SMART; SM00217; WAP; 1.
 FT NON_TER 762 762
 SQ SEQUENCE 762 AA; 86711 MW; BEB1B4EAL1CA3373 CRC64;
 Query Match 91.9%; Score 34; DB 2; Length 762;
 Best Local Similarity 85.7%; Pred. No. 1.7e+03;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ILAPVP 7
 Db 756 ILSPVP 762
 RESULT 33
 ID Q8C585 MOUSE PRELIMINARY; PRT; 768 AA.
 AC Q8C585;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Mus musculus adult male urinary bladder cDNA, RIKEN full-length
 DE enriched library, clone:9530051E16 product:weakly similar to PERQ1
 DE (Fragment).
 GN Name=Trnc15;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Urinary bladder;
 RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
 RA Carninci P., Hayashizaki Y.;
 RT "High-efficiency full-length cDNA cloning.";
 RL Meth. Enzymol. 303:19-44(1999).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Urinary bladder;
 RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
 RA Saito T., Okazaki Y., Gotojori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,

RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Urinary bladder;
 RA The FANTOM Consortium,
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 RN [4]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Urinary bladder;
 RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
 RA Carninci P., Shibata Y., Hayatsu M., Sugahara Y., Shibata K., Itoh M.,
 RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to
 RT prepare full-length cDNA libraries for rapid discovery of new genes.";
 RL Genome Res. 10:1617-1630(2000).
 RN [5]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Urinary bladder;
 RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
 RA Shibata K., Itoh M., Aizawa K., Nagasawa S., Sasaki N., Carninci P.,
 RA Konno H., Akiyama J., Nishi K., Kitsuunai T., Tashiro H., Itoh M.,
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
 RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuda S., Kawai J.,
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
 RT "RIKEN integrated sequence analysis (RISA) system-384-format
 RT sequencing pipeline with 384 multicapillary sequencer.";
 RL Genome Res. 10:1757-1771(2000).
 RN [6]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Urinary bladder;
 RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
 RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
 RA Hayashida K., Hayatsu M., Hiramoto K., Hiraoka T., Hirozane T.,
 RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
 RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
 RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
 RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
 RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
 RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
 RA Tagawa A., Takahashi F., Takaku-Akaira S., Takeda Y., Tanaka T.,
 RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK079246; BAC37586.1; -; mRNA.
 DR Ensembl; ENSMUSG0000048000; Mus musculus.
 DR MGI; MGI:2138584; Trnc15.
 DR InterPro; IPR003169; GYF.
 DR Pfam; PF02213; GYF; 1.
 DR SMART; SM00444; GYF; 1.
 DR PROSITE; PS50829; GYF; 1.
 FT NON_TER 768 768
 SQ SEQUENCE 768 AA; 87239 MW; 54FA470DF3653D57 CRC64;
 Query Match 91.9%; Score 34; DB 2; Length 768;
 Best Local Similarity 85.7%; Pred. No. 1.7e+03;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ILAPVP 7

DB 445 ILSPEVP 451
|||||

```

RESULT 34
P73172 SYNY3
AC P73172_1 SYNY3 PRELIMINARY; PRT; 924 AA.
DT 01-FEB-1997 (TRENBLrel. 02, Created)
DT 01-FEB-1997 (TRENBLrel. 02, Last sequence update)
DT 01-FEB-2004 (TRENBLrel. 26, Last annotation update)
DE Hypothetical protein.
GN OrderedLocNames=011296;
OS Synechocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX NCBI_TaxID=1148;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=97061201; PubMed=8905231;
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA Miyajima N., Hirosewa M., Sugita M., Saito M., Nakazaki N., Maruo K., Okumura S.,
RA Hoshino T., Matsumoto A., Muraki A., Wada T., Watanabe A., Yamada M., Yasuda M.,
RA Shimpo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
RA Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions.";
RL DNA Res. 3:109-136(1996).
DR EMBL; BA000022; BAA17198.1; -; Genomic_DNA.
DR PIR; S75284; S75284.
DR HSSP; Q56310; 1158.
DR GO; GO:0005737; Cytoplasm; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0000156; F:two-component response regulator activity; IEA.
DR GO; GO:0000155; F:two-component sensor molecule activity; IEA.
DR GO; GO:0006935; P:chemotaxis; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR GO; GO:0000160; P:two-component signal transduction system (p. . .); IEA.
DR InterPro; IPR003594; ATPbind_ATPase.
DR InterPro; IPR004105; H-kinase dim.
DR InterPro; IPR005467; His_kinase.
DR InterPro; IPR004358; His_kin_like_C.
DR InterPro; IPR008207; Hpt.
DR InterPro; IPR008208; Hpt N.
DR InterPro; IPR001789; Response_reg.
DR Pfam; PF01584; CheW; 1.
DR Pfam; PF02895; H-kinase_dim; 1.
DR Pfam; PF02518; HATPase_C; 1.
DR Pfam; PF01627; Hpt; 1.
DR Pfam; PF00072; BCTRLSENSOR.
DR PRINTS; PR00344; RESPONSE_REGULATORY; 1.
DR ProDom; PD003142; Hpt_N; 1.
DR ProDom; PD000039; Response_reg; 1.
DR SMART; SM00260; CheW; 1.
DR SMART; SM00387; HATPase_C; 1.
DR SMART; SM00073; HPT; 1.
DR SMART; SM00448; REC; 1.
DR PROSITE; PS01019; HIS_KIN; 1.
DR PROSITE; PS00894; HPT; 1.
DR PROSITE; PS01110; RESPONSE_REGULATORY; 1.
KW Complete proteome.
SQ SEQUENCE 924 AA; 101710 MW; 6A70A7931C814588 CRC64;

```

```

Query Match 91.9%; Score 34; DB 2; Length 924;
Best Local Similarity 71.4%; Pred. No. 2.1e+03;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 ILSPEVP 7
DB 279 VIAPPEVP 285
|||||

```

```

RESULT 35
Q520P7_MAGGR
ID Q520P7_MAGGR PRELIMINARY; PRT; 959 AA.
AC Q520P7;
DT 13-SEP-2005 (TRENBLrel. 31, Created)
DT 13-SEP-2005 (TRENBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TRENBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=MG05303.4;
OS Magnaporthe grisea 70-15.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetes incertae sedis; Magnaporthaceae; Magnaporthe.
OX NCBI_TaxID=242507;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=70-15;
RA Birren B., Nusbaum C., Abebe A., Abouelleil A., Adekoya E.,
RA Ait-zahra M., Allen N., Allen T., An P., Anderson M., Anderson S.,
RA Archchi H., Armbruster J., Bachanteang P., Baldwin J., Barry A.,
RA Bayul T., Blitshteyn B., Bloom T., Blye J., Boguslavsky L.,
RA Borowsky M., Boukhgalter B., Brunache A., Butler J., Calixte N.,
RA Calvo S., Camarata J., Campo K., Chang J., Cheshatsang Y., Citroen M.,
RA Collymore A., Considine T., Cook A., Cooke P., Corum B., Cuomo C.,
RA David R., Dawoe T., Degray S., Dodge S., Dooley K., Dorje P.,
RA Dorjee K., Dorris L., Duffey N., Dupes A., Elkins T., Engels R.,
RA Erickson J., Farina A., Faro S., Ferreira P., Fischer H.,
RA Fitzgerald M., Foley K., Gage D., Galagan J., Geatin G., Gnerre S.,
RA Glikke A., Goyette A., Graham J., Grandbois E., Gyaltsen K., Hafez N.,
RA Hagopian D., Hagos B., Hall J., Hatcher B., Heller A., Higgins H.,
RA Hanan T., Horn A., Houde N., Hughes L., Hulme W., Hubby E., Iliev I.,
RA Jaife D., Jones C., Kamal M., Kamat A., Kamveselis M., Karlsson E.,
RA Kells C., Kieu A., Kisner P., Kodira C., Kulbokas E., Labutti K.,
RA Lana D., Landers T., Leger J., Levine S., Lewis D., Lewis T.,
RA Lindblad-toh K., Liu X., Lokvitsang T., Lokvitsang Y., Lucien O.,
RA Lui A., Ma L.J., Mabbitt R., Macdonald J., Maclean C., Major J.,
RA Manning J., Marabella R., Maru K., Matthews C., Mauceli B.,
RA McCarthy M., McDonough S., McGhee T., Meldrim J., Meneus L.,
RA Meisro J., Mihalev A., Mihova T., Mikelsen T., Mlenga V., Moru K.,
RA Mozes J., Mulrain L., Munson G., Naylor J., Nazzari C., Nguyen C.,
RA Nguyen N., Nguyen T., Nicol R., Nielsen C., Nizari M., Norbu C.,
RA Norbu N., O'donnell P., Okoawo O., O'leary S., Omotoho B.,
RA O'Neill K., Oseman S., Parker S., Perrin D., Phunkhang P., Piquani B.,
RA Purcell S., Rachupka T., Ramasamy U., Rameau R., Ray V., Raymond C.,
RA Retta R., Richardson S., Rise C., Rodriguez J., Rogers J., Rogov P.,
RA Rutman M., Schupbach R., Seaman C., Settupalli S., Sharpe T.,
RA Sheridan J., Sherpa N., Shi J., Smirnov S., Smith C., Sougnez C.,
RA Spencer B., Stalker J., Stange-thomann N., Stavropoulos S.,
RA Stetson K., Stone C., Stone S., Stubbs M., Talamas J., Tchuinga P.,
RA Tenzing P., Tesfaye S., Theodore J., Thoultsang Y., Topham K.,
RA Towey S., Tsamla T., Tesmo N., Vallee D., Vassiliev H.,
RA Venkataraman V., Vinson J., Vo A., Wade C., Wang S., Wangchuk T.,
RA Wangdi T., Whittaker C., Wilkinson J., Wu Y., Wyman D., Yadav S.,
RA Yang S., Yang X., Yeager S., Yee E., Young G., Zainoun J., Zembeck L.,
RA Zimmer A., Zody M., Lander E.;
RT "The genome sequence of Magnaporthe grisea.";
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=70-15;
RA Dean R., Mitchell T., Brown D., Pan H., Thon M.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=70-15;
RA Zhu H., Blackmon B.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AACU0100810; EAA52611.1; -; Genomic_DNA.
KW Hypothetical protein.

```


SQ SEQUENCE 959 AA; 106289 MW; ADCE86F580735896 CRC64;

Query Match 91.9%; Score 34; DB 2; Length 959;
 Best Local Similarity 71.4%; Pred. No. 2.2e+03;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAPPVP 7
 : : : : :
 DB 433 VMAPVP 439

RESULT 36
 ID Q52FV3 MAGGR PRELIMINARY; PRT; 1011 AA.
 AC Q52FV3;
 DT 13-SEP-2005 (TrEMBLrel. 31, Created)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
 DE Hypothetical protein.
 GN ORFNames=MG01533.4;
 OS Magnaporthe grisea 70-15.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Sordariomycetes incertae sedis; Magnaporthaceae; Magnaporthae.
 OX NCBI_TaxID=242507;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=70-15;
 RA Birren B., Nusbaum C., Abebe A., Abouelleil A., Adekoya E.,
 RA Ait-zahra M., Allen N., Allen T., An P., Anderson M., Anderson S.,
 RA Archachi H., Armbruster J., Bachantang P., Baldwin J., Barry A.,
 RA Bayul T., Blitsherev B., Bloom T., Blye J., Boguslavskiy L.,
 RA Borowsky M., Boukhgalter B., Brunache A., Butler J., Calixte N.,
 RA Calvo S., Camarata J., Campo K., Chang J., Cheshnatsang Y., Citroen M.,
 RA Collymore A., Considine T., Cook A., Cooke P., Corum B., Cuomo C.,
 RA David R., Dawot L., Duffey N., Dupes A., Elkins T., Engels R.,
 RA Dorjee K., Dorris L., Duffey N., Duffey S., Dooley K., Dorje P.,
 RA Erickson J., Farina A., Faro S., Ferreira P., Fischer H.,
 RA Fitzgerald M., Foley K., Gage D., Galagan J., Gearin G., Gnerre S.,
 RA Gnrke A., Goyette A., Graham J., Grandbois E., Gyaltsang Y., Hafez N.,
 RA Hagopian D., Hagos B., Hall J., Hatcher B., Heller A., Higgins H.,
 RA Honan T., Horn A., Houde N., Hughes L., Hulme W., Husby E., Iliev I.,
 RA Jaffe D., Jones C., Kamal M., Kamat A., Kamysseelis M., Karlsson E.,
 RA Kells C., Kieu A., Kisner P., Kodira C., Kulbokas E., Labutti K.,
 RA Lama D., Landers T., Leger J., Levine S., Lewis D., Lewis T.,
 RA Lindblad-toh K., Liu X., Lokyitsang T., Lokyitsang Y., Lucien O.,
 RA Lui A., Ma L.J., Mabbitt R., Macdonald J., Maclean C., Major J.,
 RA Manning J., Marabella R., Maru K., Matthews C., Mauceli E.,
 RA McCarthy M., McDonough S., Mcghee T., Meldrim J., Meneus L.,
 RA Mesirov J., Mihalev A., Mihova T., Mikkelsen T., Mienga V., Moru K.,
 RA Mozes J., Mullrain L., Munson G., Naylor J., Neues C., Nguyen C.,
 RA Nguyen N., Nguyen T., Nicol R., Nielsen C., Nizzari M., Norbu C.,
 RA Norbu N., O'donnell P., Okosawo O., O'leary S., Omotosho B.,
 RA O'Neill K., Oman S., Parker S., Perrin D., Phunkhang P., Pigani B.,
 RA Purcell S., Rachupka T., Ramasamy U., Rameau R., Ray V., Raymond C.,
 RA Retta R., Richardson S., Rise C., Rodriguez J., Rogers J., Rogov P.,
 RA Rutman M., Schnupbach R., Seaman C., Settippalli S., Sharpe T.,
 RA Sheridan J., Sherpa N., Shi J., Smirnov S., Smith C., Sougnez C.,
 RA Spencer B., Stalker J., Stange-thomann N., Stavropoulos S.,
 RA Stetson K., Stone C., Stone S., Stubbs M., Talamas J., Tchuinga P.,
 RA Tenxing P., Tesfaye S., Theodore J., Thoultsang Y., Topham K.,
 RA Towey S., Tsamila T., Tsomo N., Vallee D., Vassiliev H.,
 RA Venkataraman V., Vinson J., Vo A., Wade C., Wang S., Wangchuk T.,
 RA Wangdi T., Whittaker C., Wilkinson J., Wu Y., Wyman D., Yadav S.,
 RA Yang S., Yang X., Yeager S., Yee E., Young G., Zainoun J., Zembeck L.,
 RA Zimmer A., Zody M., Lander E.;
 FT "The genome sequence of Magnaporthe grisea";
 RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
 RN [2]

RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=70-15;
 RA Dean R., Mitchell T., Brown D., Pan H., Thon M.;
 RA Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
 RN [3]

RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=70-15;
 RA Zhu H., Blackmon B.;
 RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
 CC -!- CAUTION: The sequence shown here is derived from an
 EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 KW EMBL: AACU0100067; EAA55882.1; -; Genomic_DNA.
 SQ SEQUENCE 1011 AA; 107576 MW; 81A6BBA934D950AB CRC64;

Query Match 91.9%; Score 34; DB 2; Length 1011;
 Best Local Similarity 85.7%; Pred. No. 2.3e+03;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAPPVP 7
 : : : : :
 DB 139 ILAPPVP 145

RESULT 37
 Q80TVI MOUSE
 ID Q80TVI MOUSE PRELIMINARY; PRT; 1048 AA.
 AC Q80TVI;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE MKIAA0642 protein (Fragment).
 GN Name=Trnc15; Synonyms=mkIAA0642;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Brain;
 RX MEDLINE=22579291; PubMed=12693553;
 RA Okazaki N., Kikuno R., Ohara R., Inamoto S., Aizawa H., Yuasa S.,
 RA Nakajima D., Nagase T., Ohara O., Koga H.;
 RT "Prediction of the coding sequences of mouse homologues of KIAA gene:
 RT II. The complete nucleotide sequences of 400 mouse KIAA-homologous
 RT cDNAs identified by screening of terminal sequences of cDNA clones
 RT randomly sampled from size-fractionated libraries.";
 RL DNA Res. 10:35-48(2003).
 DR EMBL: AK122337; BAC65619.1; -; mRNA.
 DR HSSP; Q05195; INLM.
 DR Ensembl; ENSMUSG00000048000; Mus musculus.
 DR MGI; MGI:2138584; Trnc15.
 DR InterPro; IPR003169; GYF.
 DR Pfam; PF02213; GYF; 1.
 DR SMART; SM00444; GYF; 1.
 DR PROSITE; PS50829; GYF; 1.
 FT NON TER 1
 SQ SEQUENCE 1048 AA; 121839 MW; 113C516E81E98A4C CRC64;

Query Match 91.9%; Score 34; DB 2; Length 1048;
 Best Local Similarity 85.7%; Pred. No. 2.4e+03;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAPPVP 7
 : : : : :
 DB 208 ILSPVP 214

RESULT 38
 Q6Y7W8 MOUSE
 ID Q6Y7W8 MOUSE PRELIMINARY; PRT; 1291 AA.
 AC Q6Y7W8;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Grb10 interacting GYF protein 2.

```

GN Name=Trnc15;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Liver;
RX MEDLINE=22803289; PubMed=12771153; DOI=10.1074/jbc.M211572200;
RA Giovannone B., Lee E., Laviola L., Giorgio F., Cleveland K.A.,
RA Smith R.J.;
RT "Two novel proteins that are linked to insulin-like growth factor
RT (IGF-I) receptors by the Grb10 adapter and modulate IGF-I signaling.";
RL J. Biol. Chem. 278:31564-31573(2003).
DR EMBL; AY176043; AA046887.1; -; mRNA.
DR MGI; MGI:2138584; Trnc15.
DR InterPro; IPR003169; GVF.
DR Pfam; PF02213; GVF; 1.
DR SMART; SM00444; GVF; 1.
DR PROSITE; PS50829; GVF; 1.
SQ SEQUENCE 1291 AA; 149173 MW; C0B423D99DF0BC5E CRC64;

Query Match 91.9%; Score 34; DB 2; Length 1291;
Best Local Similarity 85.7%; Pred. No. 3e+03; 0; Indels 0; Gaps 0;
Matches 6; Conservative 1; Mismatches 0;

Qy 1 ILAPPVP 7
Db 451 ILSPVP 457

RESULT 39
Q85726_9CAUD
ID Q85726_9CAUD PRELIMINARY; PRT; 111 AA.
AC Q85726_
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DE 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Gp15.
OS Mycobacteriophage Cjw1.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae.
OX NCBI_TaxID=205863;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22592660; PubMed=12705866; DOI=10.1016/S0092-8674(03)00233-2;
RA Pedulla M.L., Ford M.E., Houtz J.M., Karthikeyan T., Wadsworth C.,
RA Lewis J.A., Jacobs-Sera D., Falbo J., Gross J., Pannunzio N.R.,
RA Brucker W., Kumar V., Kandasamy J., Keenan L., Bardarov S.,
RA Kriakov J., Lawrence J.G., Jacobs W.R. Jr., Hendrix R.W.,
RA Hatfull G.F.;
RT "Origins of highly mosaic mycobacteriophage genomes.";
RL Cell 113:171-182(2003).
DR EMBL; AY129331; AA01630.1; -; Genomic DNA.
SQ SEQUENCE 111 AA; 12237 MW; A6B87B381477527 CRC64;

Query Match 89.2%; Score 33; DB 2; Length 111;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LAPPVP 7
Db 62 LAPPVP 67

RESULT 40
Y124_LISIN
ID Y124_LISIN STANDARD; PRT; 113 AA.
AC Q925X3;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Hypothetical protein Lin0124/Lin2378/Lin2564.

```

```

GN OrderedLocusNames=lin0124;
GN and
GN OrderedLocusNames=lin2378;
GN and
GN OrderedLocusNames=lin2564;
OS Listeria innocua.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX NCBI_TaxID=1642;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=CLIP 11262 / Serovar 6a;
RX MEDLINE=21537279; PubMed=11679669; DOI=10.1126/science.1063447;
RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Anand A.,
RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty P.,
RA Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapkat G.,
RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordsiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schluster T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
RT "Comparative genomics of Listeria species.";
RL Science 294:849-852(2001).
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; AL596163; CAC95357.1; -; Genomic DNA.
DR EMBL; AL596172; CAC97605.1; -; Genomic DNA.
DR EMBL; AL596172; CAC97791.1; -; Genomic DNA.
DR PIR; AE1448; AE1448.
DR PIR; AE1729; AE1729.
DR PIR; AG1752; AG1752.
DR ListList; LIN00124; -.
DR ListList; LIN02378; -.
DR ListList; LIN02564; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 113 AA; 12527 MW; D2B7F93D13BD4E32 CRC64;

Query Match 89.2%; Score 33; DB 1; Length 113;
Best Local Similarity 71.4%; Pred. No. 3.4e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILAPPVP 7
Db 36 LIAPPVP 42

RESULT 41
Q8DKV4_SYNEL
ID Q8DKV4_SYNEL PRELIMINARY; PRT; 129 AA.
AC Q8DKV4;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Dihydroneopterin adolase.
GN OrderedLocusNames=tl10747;
OS Synecococcus elongatus (Thermosynechococcus elongatus).
OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
OX NCBI_TaxID=32046;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=BP-1;
RX MEDLINE=22225144; PubMed=12240834;
RA Nakamura Y., Kaneko T., Sato S., Ikeuchi M., Katoh H., Sasamoto S.,
RA Watanabe A., Iriguchi M., Kawashima K., Kimura T., Kishida Y.,
RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Nakazaki N.,
RA Shimo S., Sugimoto M., Takeuchi C., Yamada M., Tabata S.;

```

```

RT "Complete genome structure of the thermophilic cyanobacterium
RT Thermosynechococcus elongatus BP-1.";
RL DNA Res. 9:123-130(2002).
DR EMBL; BA000039; BAC08298.1; -; Genomic_DNA.
DR HSSP; P56740; 1DHN.
DR GO; GO:0004150; P:dihydroneopterin aldolase activity; IEA.
DR GO; GO:0005760; P:folic acid and derivative metabolism; IEA.
DR InterPro; IPR006157; F0LB_fam.
DR InterPro; IPR006156; F0LB_fam.
DR Pfam; PF02152; F0LB; 1.
DR TIGRFAMs; TIGR00525; f0LB; 1.
DR TIGRFAMs; TIGR00526; f0LB_dom; 1.
KW Complete proteome.
SQ SEQUENCE 129 AA; 14567 MW; 689CC2E1C5BECD20 CRC64;

Query Match 89.2%; Score 33; DB 2; Length 129;
Best Local Similarity 100.0%; Pred. No. 3.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LAPPVP 7
DB 106 LAPPVP 111

RESULT 42
Q9RGAL 9RHIZ PRELIMINARY; PRT; 147 AA.
AC Q9RGAL;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2004 (TrEMBLrel. 26, Last annotation update)
DE Tioflf6 protein.
GN Name=tioflf6;
OS Agrobacterium tumefaciens.
OG Plasmid pTi-SAKURA.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Agrobacterium.
OX NCBI_TaxID=356;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MAFP301001;
RA Hattori Y., Suzuki K., Ohta N., Uraji M., Katoh A., Yoshida K.;
RT "Genome structure of pTi-SAKURA (I): Strategy for DNA sequencing of a
RT Japanese cherry-Ti plasmid.";
RL Nucleic Acids Symp. Ser. 37:159-160(1998).
[2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MAFP301001;
RA Ohta N., Suzuki K., Hattori Y., Uraji M., Katoh A., Yoshida K.;
RT "Genome structure of pTi-SAKURA (III): Characteristics of T-DNA.";
RL Nucleic Acids Symp. Ser. 39:185-186(1998).
[3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MAFP301001;
RA Uraji M., Suzuki K., Ohta N., Hattori Y., Katoh A., Yoshida K.;
RT "Genome structure of pTi-SAKURA (IV): Characteristics of tra region.";
RL Nucleic Acids Symp. Ser. 39:187-188(1998).
[4]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MAFP301001;
RA Hattori Y., Suzuki K., Ohta N., Uraji M., Katoh A., Yoshida K.;
RT "Genome structure of pTi-SAKURA (V): Complete nucleotide sequence of
RT plasmid pTi-SAKURA's vir region in Agrobacterium tumefaciens.";
RL Nucleic Acids Symp. Ser. 39:265-266(1998).
[5]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MAFP301001;
RX MEDLINE=20184752; PubMed=10721727; DOI=10.1016/S0378-1119(99)00502-8;
RA Suzuki K., Hattori Y., Uraji M., Ohta N., Iwata K., Murata K.,
RA Katoh A., Yoshida K.;
RT "Complete nucleotide sequence of a plant tumor-inducing Ti plasmid.";
RL Gene 242:331-336(2000).
[6]

RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MAFP301001;
RX MEDLINE=98193120; PubMed=9524202; DOI=10.1016/S0167-4781(97)00182-6;
RA Suzuki K., Ohta N., Hattori Y., Uraji M., Katoh A., Yoshida K.;
RT "Novel structural difference between nopaline- and octopine- type trbJ
RT gene: construction of genetic and physical map and sequencing of
RT trbJ/traI and rep gene clusters of a new Ti plasmid pTi-SAKURA.";
RL Biochim. Biophys. Acta 1396:1-7(1998).
DR EMBL; AB016260; BAA87801.1; -; Genomic_DNA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0004803; P:transposase activity; IEA.
DR GO; GO:0006313; P:DNA transposition; IEA.
DR InterPro; IPR002514; Transposase_8.
DR Pfam; PF01527; Transposase_8; 1.
KW Plasmid.
SQ SEQUENCE 147 AA; 15964 MW; 7B72319E1B9391EE CRC64;

Query Match 89.2%; Score 33; DB 2; Length 147;
Best Local Similarity 71.4%; Pred. No. 4.5e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAPPVP 7
DB 102 VLAPPVP 108

RESULT 43
Q9K3W9 STRCO PRELIMINARY; PRT; 164 AA.
AC Q9K3W9;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative marR-family regulatory protein.
GN OrderedLocustNames=SCO4375; ORFNames=SCD10.07;
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953; DOI=10.1038/417141a;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L.D., Oliver K., O'Neill S.,
RA Rabinowitz E., Rajandream M.A., Rutherford K.M., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorek A., Woodward J.R., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RL Nature 417:141-147(2002).
DR EMBL; AL939119; CAB95886.1; -; Genomic_DNA.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR000835; HTH_MarR.
DR InterPro; IPR011991; Wing_hlx_DNA_bd.
DR Pfam; PF01047; MarR; 1.
DR PRINTS; PR00598; HTHMARR.
DR SMART; SM00347; HTH_MARR; 1.
DR PROSITE; PS50995; HTH_MARR_2; 1.
KW Complete proteome; DNA-binding; Transcription;
KW Transcription regulation.
SQ SEQUENCE 164 AA; 18067 MW; 244067D79F05FEDB CRC64;

Query Match 89.2%; Score 33; DB 2; Length 164;
Best Local Similarity 100.0%; Pred. No. 5.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      2 LAPPVP 7
Db      149 LAPPVP 154

RESULT 44
Q629K3_ORYZA
ID      Q629K3_ORYZA PRELIMINARY;      PRT;      177 AA.
AC      Q629K3
DT      05-JUL-2004 (TREMBLrel. 27, Created)
DT      05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DT      05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE      Hypothetical protein P0475A09.28.
GN      Name=P0475A09.28;
OS      Oryza sativa (japonica cultivar-group).
OC      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC      Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC      Ehrhartoideae; Oryzaceae; Oryza.
OX      NCBI_TaxID=39947;
RN      [1]
RP      NUCLEOTIDE SEQUENCE.
RA      Sasaki T., Matsumoto T., Yamamoto K.;
RT      "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 8, PAC
RT      clone:P0475A09."
RL      Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR      EMBL; AP004696; BAD01331.1; -; Genomic_DNA.
DR      Gramene; Q629K3; -.
KW      Hypothetical protein.
SQ      SEQUENCE 177 AA; 18335 MW; 426AB8B8A5E3952BF CRC64;

Query Match      89.2%; Score 33; DB 2; Length 177;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 LAPPVP 7
Db      106 LAPPVP 111

RESULT 45
Q73TA7_MYCPA
ID      Q73TA7_MYCPA PRELIMINARY;      PRT;      183 AA.
AC      Q73TA7;
DT      05-JUL-2004 (TREMBLrel. 27, Created)
DT      05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DT      05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE      Hypothetical protein.
GN      OrderedLocustNames=MAP3811;
OS      Mycobacterium paratuberculosis.
OC      Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC      Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC      Mycobacterium avium complex (MAC).
OX      NCBI_TaxID=1770;
RN      [1]
RP      NUCLEOTIDE SEQUENCE.
RC      STRAIN=K10;
RA      Li L., Bannantine J., Zhang Q., Anonin A., Alt D., Kapur V.;
RL      Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
KW      Complete proteome; Hypothetical protein.
SQ      SEQUENCE 183 AA; 16640 MW; D7BE0F54CE6D04E1 CRC64;

Query Match      89.2%; Score 33; DB 2; Length 183;
Best Local Similarity 100.0%; Pred. No. 5.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 LAPPVP 7
Db      118 LAPPVP 123

RESULT 46
HLH4_CAEEL

```

```

ID      HLH4_CAEEL      STANDARD;      PRT;      205 AA.
AC      P34555;
DT      01-FEB-1994 (Rel. 28, Created)
DT      01-FEB-1994 (Rel. 28, Last sequence update)
DT      13-SEP-2005 (Rel. 48, Last annotation update)
DE      Helix-loop-helix protein 4.
GN      Name=hlh-4; ORFNames=T05G5.2;
OS      Caenorhabditis elegans.
OC      Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC      Rhabditidae; Peloderinae; Caenorhabditis.
OX      NCBI_TaxID=6239;
RN      [1]
RP      NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC      STRAIN=Bristol N2;
RX      MEDLINE=94150718; PubMed=7906398; DOI=10.1038/368032a0;
RA      Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA      Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA      Craxton M., Dear S., Du Z., Durbin R., Pavello A., Fraser A.,
RA      Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
RA      Johnston L., Jones M., Kershaw J., Kirsten J., Laister N.,
RA      Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
RA      Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,
RA      Sims M., Smaildon N., Smith A., Smith M., Sonnhammer E., Staden R.,
RA      Suleston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
RA      Waterston R., Watson A., Weinstein L., Wilkinson-Sproat J.,
RA      Woudman P.;
RT      "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT      elegans."
RL      Nature 368:32-38(1994).
RN      [2]
RP      NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC      STRAIN=Bristol N2;
RX      MEDLINE=99069613; PubMed=9851916;
RG      The C. elegans sequencing consortium;
RT      "Genome sequence of the nematode C. elegans: a platform for
RT      investigating biology."
RL      Science 282:2012-2018(1998).
CC      -|- SUBCELLULAR LOCATION: Nuclear (Potential).
CC      -|- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
CC      -----
CC      This Swiss-Prot entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use as long as its content is in no way modified and this statement is not
CC      removed.
CC      -----
DR      EMBL; Z27079; CAA81589.1; -; Genomic_DNA.
DR      PIR; S41002; S41002.
DR      Ensembl; T05G5.2; Caenorhabditis elegans.
DR      WormBase; WBGene00001951; hlh-4.
DR      WormPep; T05G5.2; CE00314.
DR      InterPro; IPR001092; HLH_basic.
DR      Pfam; PF00010; HLH; 1.
DR      SMART; SM00353; HLH; 1.
DR      PROSITE; PS00888; HLH; 1.
KW      Complete proteome; Hypothetical protein; Nuclear protein.
FT      DOMAIN 38..89 Helix-loop-helix motif (by similarity).
SQ      SEQUENCE 205 AA; 22692 MW; 577E59E04268FA1C CRC64;

Query Match      89.2%; Score 33; DB 1; Length 205;
Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 LAPPVP 7
Db      139 LAPPVP 144

RESULT 47
P97177_RHOSH
ID      P97177_RHOSH PRELIMINARY;      PRT;      226 AA.
AC      P97177;
DT      01-MAY-1997 (TREMBLrel. 03, Created)

```

DT 01-MAY-1997 (TReMBLrel. 03, Last sequence update)
 DT 01-MAY-2004 (TReMBLrel. 26, Last annotation update)
 DE Transmembrane protein.
 GN Name=dmsB;
 OS Rhodobacter sphaeroides (Rhodospseudomonas sphaeroides).
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales;
 OC Rhodobacteraceae; Rhodobacter.
 OX NCBI_TaxID=1063;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=97107628; PubMed=8950368; DOI=10.1016/S0005-2728(96)00101-6;
 RA Ujiiye T., Yamamoto I., Nakama H., Okubo A., Yamazaki S., Satoh T.;
 RT "Nucleotide sequence of the genes, encoding the pentaheme cytochrome
 (dmsC) and the transmembrane protein (dmsB), involved in dimethyl
 sulfoxide respiration from Rhodobacter sphaeroides f. sp.
 denitrificans";
 RL Biochim. Biophys. Acta 1277:1-5(1996).
 DR EMBL; D82820; BAA11582.1; -; Genomic_DNA.
 DR PIR; PC4101;
 DR GO; GO:0005737; Cytoplasm; IEA.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0051085; P:chaperone cofactor dependent protein folding; IEA.
 DR InterPro; IPR010395; TorD.
 DR Pfam; PF06192; TorD; 1.
 KW Transmembrane.
 SQ SEQUENCE 226 AA; 23551 MW; F821A37FC21B6A3F CRC64;

 Query Match 89.2%; Score 33; DB 2; Length 226;
 Best Local Similarity 100.0%; Pred. No. 7.1e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 2 LAPPVP 7
 DB |||||
 3 LAPPVP 8

 RESULT 48
 ID Q5FC67 CAEBL PRELIMINARY; PRT; 228 AA.
 AC Q5FC67;
 DT 10-MAY-2005 (TReMBLrel. 30, Created)
 DT 10-MAY-2005 (TReMBLrel. 30, Last sequence update)
 DT 10-MAY-2005 (TReMBLrel. 30, Last annotation update)
 DE Hypothetical protein T06E4.12;
 GN ORFNames=T06E4.12;
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=Bristol N2;
 RX MEDLINE=99069613; PubMed=9851916;
 RG The C. elegans sequencing consortium;
 RT "Genome sequence of the nematode C. elegans: a platform for
 investigating biology.";
 RL Science 282:2012-2018(1998).
 DR EMBL; Z70756; CAI46587.1; -; Genomic DNA.
 DR Ensembl; T06E4.12; Caenorhabditis elegans.
 DR WormBase; WBGene0044011; T06E4.12.
 DR WormPep; T06E4.12; CE37848.
 DR GO; GO:0005199; F:structural constituent of cell wall; IEA.
 DR InterPro; IPR003882; Pistil extensin.
 DR PRINTS; PR01218; PSTLEXTENSIN.
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 228 AA; 22221 MW; FIDEB689484B8913 CRC64;

 Query Match 89.2%; Score 33; DB 2; Length 228;
 Best Local Similarity 100.0%; Pred. No. 7.2e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 2 LAPPVP 7
 DB |||||

Db 112 LAPPVP 117

 RESULT 49
 ID Q52WY3 BRARE PRELIMINARY; PRT; 235 AA.
 AC Q52WY3;
 DT 13-SEP-2005 (TReMBLrel. 31, Created)
 DT 13-SEP-2005 (TReMBLrel. 31, Last sequence update)
 DT 13-SEP-2005 (TReMBLrel. 31, Last annotation update)
 DE Ras-like protein RhoB (Fragment).
 OS Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 OX NCBI_TaxID=7955;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX PubMed=15894457; DOI=10.1016/j.ygeno.2005.03.010;
 RA Salas-Vidal E., Meijer A.H., Cheng X., Spalink H.P.;
 RT "Genomic annotation and expression analysis of the zebrafish Rho small
 GTPase family during development and bacterial infection.";
 RL Genomics 86:25-37(2005).
 DR EMBL; AY865565; AAX20137.1; -; mRNA.
 DR GO; GO:0005525; F:GTP binding; IEA.
 DR GO; GO:0006886; P:intracellular protein transport; IEA.
 DR GO; GO:0007264; P:small GTPase mediated signal transduction; IEA.
 DR InterPro; IPR003579; GTPase Ras.
 DR InterPro; IPR003577; GTPase Ras.
 DR InterPro; IPR003578; GTPase_Rho.
 DR InterPro; IPR002041; RAN.
 DR InterPro; IPR001806; Ras trnsfrmg.
 DR InterPro; IPR005225; Small_GTP.
 DR Pfam; PF00071; Ras; 1.
 DR PRINTS; PR00449; RASTRNSFRMNG.
 DR SMART; SM00175; RAB; 1.
 DR SMART; SM00176; RAN; 1.
 DR SMART; SM00173; RAS; 1.
 DR SMART; SM00174; RHO; 1.
 DR TIGRFAMs; TIGR00231; small GTP; 1.
 KW GTP-binding; Nucleotide-binding.
 FT NON_TER 235 235
 SQ SEQUENCE 235 AA; 26491 MW; B79EACC9FC0EF5D4 CRC64;

 Query Match 89.2%; Score 33; DB 2; Length 235;
 Best Local Similarity 71.4%; Pred. No. 7.4e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

 QY 1 ILAPPVP 7
 DB ::|||
 6 LMAPPVP 12

 RESULT 50
 ID Q89E19 BRAJA PRELIMINARY; PRT; 236 AA.
 AC Q89E19;
 DT 01-JUN-2003 (TReMBLrel. 24, Created)
 DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
 DE Bll7096 protein.
 GN OrderedLocusNames=bll7096;
 OS Bradyrhizobium japonicum.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Bradyrhizobiaceae; Bradyrhizobium.
 OX NCBI_TaxID=375;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=USDA 110;
 RX MEDLINE=22484998; PubMed=12597275;
 RA Kaneko T., Nakamura Y., Sato S., Minamieawa K., Uchiumi T.,
 RA Sasamoto S., Watanabe A., Ideawa K., Iriguchi M., Kawashima K.,
 RA Kohara M., Matsumoto M., Shimpo S., Tsuruoka H., Wada T., Yamada M.,

```

RA Tabata S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RL Bradyrhizobium japonicum USDA110.";
RL DNA Res. 9:189-197(2002).
DR EMBL; BA000040; BACS2361.1; -; Genomic_DNA.
DR InterPro; IPR007400; DUF453.
DR Pfam; PF04303; DUF453; 1.
KW Complete proteome.
SQ SEQUENCE 236 AA; 24727 MW; BB79EEB04F48421B CRC64;

Query Match      89.2%; Score 33; DB 2; Length 236;
Best Local Similarity 100.0%; Pred.No. 7.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 LAPPVP 7
DB      183 LAPPVP 188

Search completed: April 6, 2006, 09:39:26
Job time : 112.105 secs

```

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: April 6, 2006, 09:22:55 ; Search time 78.2895 Seconds
(without alignments)
39.286 Million cell updates/sec

Title: US-10-632-388-305

Perfect score: 45

Sequence: 1 RPQPPPP 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Database : A_Geneseq_21.*

- 1: Geneseqp1980s.*
- 2: Geneseqp1990s.*
- 3: Geneseqp2000s.*
- 4: Geneseqp2001s.*
- 5: Geneseqp2002s.*
- 6: Geneseqp2003as.*
- 7: Geneseqp2003bs.*
- 8: Geneseqp2004s.*
- 9: Geneseqp2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	45	100.0	7	3	AAB17249
2	45	100.0	7	5	ABB73242
3	45	100.0	7	7	ADJ73396
4	45	100.0	7	8	ADJ53030
5	45	100.0	7	8	ADJ51991
6	45	100.0	13	2	AAW11113
7	45	100.0	31	2	AAW25511
8	45	100.0	122	5	ADJ45407
9	45	100.0	173	5	ABB97921
10	45	100.0	249	7	ABO72442
11	45	100.0	276	8	ADP29798
12	45	100.0	288	3	AAG13899
13	45	100.0	314	9	ABM97421
14	45	100.0	322	3	AAG13898
15	45	100.0	340	3	AAG06370
16	45	100.0	398	3	AAG06370
17	45	100.0	398	8	ADN73283
18	45	100.0	412	3	AAG06369
19	45	100.0	432	2	AAW72182
20	45	100.0	446	4	ABB70063
21	45	100.0	497	2	AAW72030
22	45	100.0	497	2	AAW72132
23	45	100.0	527	2	AAI29515
24	45	100.0	527	3	AAB44492

25	45	100.0	527	4	AAE13805
26	45	100.0	527	7	ADD6524
27	45	100.0	527	7	ADG87778
28	45	100.0	668	5	AAU10545
29	45	100.0	669	9	ADK05987
30	45	100.0	704	5	AAU10544
31	45	100.0	704	7	ADE57712
32	45	100.0	705	4	AM339328
33	45	100.0	705	9	ADK05985
34	45	100.0	714	4	AAU11114
35	45	100.0	735	4	ABG26213
36	45	100.0	812	4	ABG30148
37	45	100.0	874	4	ABG02079
38	45	100.0	1074	4	ABG00401
39	45	100.0	1413	5	AAE21729
40	45	100.0	1460	7	ADC59465
41	45	100.0	1460	8	ADU04632
42	45	100.0	1734	6	ABR42259
43	45	100.0	1734	7	ABM78986
44	45	100.0	1734	8	ADG20497
45	45	100.0	1734	8	ADQ15178
46	45	100.0	1798	5	AAU49177
47	45	100.0	1798	5	AAE16277
48	45	100.0	1798	8	ADJ96613
49	45	100.0	1805	8	ADK91705
50	45	100.0	4952	5	ADH47759
51	45	100.0	4952	7	ADP68294
52	45	100.0	4952	8	ADL25642
53	45	100.0	4981	8	ADN05339
54	45	100.0	5159	5	ADH48828
55	42	93.3	129	8	ADK94859
56	42	93.3	216	4	ABE65033
57	42	93.3	279	8	ADR10339
58	42	93.3	424	7	ABM88266
59	42	93.3	440	7	ABM90225
60	42	93.3	549	8	ADI45327
61	42	93.3	626	4	AAU40777
62	42	93.3	1222	3	AAK36453
63	42	93.3	1257	3	AAK36452
64	42	93.3	1264	8	ADQ66678
65	42	93.3	1275	3	AAK36451
66	42	93.3	1278	2	AAW88445
67	42	93.3	1278	3	AAK42983
68	42	93.3	1278	8	ADQ39879
69	42	93.3	1278	8	ADU06723
70	42	93.3	1639	4	ABB59281
71	42	93.3	1745	5	ABB57253
72	41	91.1	68	4	ABG11143
73	41	91.1	68	5	AAE26653
74	41	91.1	68	5	ABP64470
75	41	91.1	68	5	ABP72613
76	41	91.1	70	8	ADT93238
77	41	91.1	79	4	AAK69616
78	41	91.1	87	5	ABG30880
79	41	91.1	87	6	ABP72612
80	41	91.1	89	7	ABM89363
81	41	91.1	89	9	ADU77066
82	41	91.1	91	8	ADH09438
83	41	91.1	91	8	ADU82410
84	41	91.1	92	8	ADU82414
85	41	91.1	92	8	ADU82414
86	41	91.1	92	8	ADU82417
87	41	91.1	92	8	ADU82419
88	41	91.1	92	8	ADU82416
89	41	91.1	92	8	ADU82415
90	41	91.1	92	8	ADU82418
91	41	91.1	92	8	ADU82420
92	41	91.1	98	4	AAU62338
93	41	91.1	98	6	ABM58857
94	41	91.1	108	2	AAW95071
95	41	91.1	108	2	AAW95076
96	41	91.1	108	2	ADJ48368
97	41	91.1	112	8	ADK67103

98	41	91.1	114	8	ADS15397	Adsl5397 HSV-1 pol	171	41	91.1	724	8	ABO84902	ABO84902 Human can
99	41	91.1	121	9	ADV77067	Adv77067 Huntingto	172	41	91.1	726	7	ABO71929	ABO71929 Pseudomon
100	41	91.1	122	3	AGV33269	Agv33269 Zeta maye	173	41	91.1	733	4	ABB71995	ABB71995 Drosophil
101	41	91.1	142	4	ABG18931	Abg18931 Novel hum	174	41	91.1	736	5	ABB57361	ABB57361 Mouse isc
102	41	91.1	144	7	ADF60189	Adf60189 Human con	175	41	91.1	745	8	ABO84899	ABO84899 Murine cod
103	41	91.1	142	7	ABO77746	AbO77746 Pseudomon	176	41	91.1	780	8	ADI82549	ADI82549 Human mod
104	41	91.1	157	7	ABM88176	Abm88176 Rice abio	177	41	91.1	784	3	AAH11694	AAH11694 Human sec
105	41	91.1	164	8	ADU82409	Adu82409 Mutant Un	178	41	91.1	787	3	AAH11693	AAH11693 Human sec
106	41	91.1	171	2	AAW99022	Aaw99022 Human hun	179	41	91.1	825	7	ADJ70479	ADJ70479 Human hea
107	41	91.1	171	5	AAE26650	Aae26650 Human hun	180	41	91.1	826	3	ADM48205	ADM48205 Polypteti
108	41	91.1	194	8	ADU01204	Adu01204 Human pro	181	41	91.1	837	3	RAY94903	RAY94903 Human sec
109	41	91.1	194	8	ADU15586	Adu15586 Novel hum	182	41	91.1	900	3	AAH42321	AAH42321 Human ORF
110	41	91.1	197	8	ADK68833	Adk68833 Plant ful	183	41	91.1	945	2	AAV06119	AAV06119 Human CII
111	41	91.1	216	7	ABM89871	Abm89871 Rice abio	184	41	91.1	972	5	ABE92750	ABE92750 Herbicida
112	41	91.1	218	8	ADK87616	Adk87616 Plant ful	185	41	91.1	978	9	ADU68354	ADU68354 Human kio
113	41	91.1	226	4	ABH71956	Abh71956 Drosophil	186	41	91.1	979	2	AAAR08338	AAAR08338 Complete
114	41	91.1	237	8	ADY13056	Ady13056 Plant ful	187	41	91.1	1012	2	AAW63671	AAW63671 Polypteti
115	41	91.1	248	8	ADQ26769	Adq26769 Human her	188	41	91.1	1012	4	AAW73614	AAW73614 Human kio
116	41	91.1	248	8	ADQ26768	Adq26768 Human her	189	41	91.1	1012	7	ADJ69610	ADJ69610 Human hea
117	41	91.1	248	8	ADS15399	Ads15399 HSV-1 pol	190	41	91.1	1012	8	ADP43058	ADP43058 Human kio
118	41	91.1	255	6	ABP56850	Abp56850 Zebrafish	191	41	91.1	1012	8	AQD39809	AQD39809 Human myo
119	41	91.1	264	7	ABM89291	Abm89291 Rice abio	192	41	91.1	1012	9	ADU68353	ADU68353 Human kio
120	41	91.1	309	8	ADY04216	Ady04216 Plant ful	193	41	91.1	1048	4	ABG20939	ABG20939 Novel hum
121	41	91.1	315	9	ABE09208	AbE09208 Huntingti	194	41	91.1	1048	8	ADS11844	ADS11844 Human the
122	41	91.1	317	4	AAU67582	Aau67582 Propionib	195	41	91.1	1109	7	ADC59312	ADC59312 Human pol
123	41	91.1	317	6	ABM64101	Abm64101 Propionib	196	41	91.1	1217	4	ABG09876	ABG09876 Novel hum
124	41	91.1	322	9	ADV77064	Adv77064 Huntingto	197	41	91.1	1239	4	ABG09877	ABG09877 Novel hum
125	41	91.1	330	7	ABO69642	AbO69642 Pseudomon	198	41	91.1	1277	9	ABEB32343	ABEB32343 Human pro
126	41	91.1	331	8	ADT93240	Adt93240 Human Hun	199	41	91.1	1278	7	ABM85431	ABM85431 Human pro
127	41	91.1	334	6	ABM64819	Abm64819 Propionib	200	41	91.1	1313	7	ADJ70374	ADJ70374 Human hea
128	41	91.1	334	7	ADM05556	Adm05556 Human pro	201	41	91.1	1324	8	ADJ66660	ADJ66660 IRS2 prot
129	41	91.1	357	7	ADJ80172	Adj80172 Novel hum	202	41	91.1	1324	8	ADN98330	ADN98330 Human ins
130	41	91.1	361	8	ADI42062	Adi42062 Plant tra	203	41	91.1	1324	8	ADR14675	ADR14675 Human NF-
131	41	91.1	361	8	ADO02607	Ado02607 Thalecres	204	41	91.1	1324	9	ADU87428	ADU87428 Human IRS
132	41	91.1	380	8	ADY24834	Ady24834 Plant ful	205	41	91.1	1324	9	ABEB32315	ABEB32315 Human pro
133	41	91.1	384	7	ADY04877	Ady04877 Plant ful	206	41	91.1	1338	8	ADO19685	ADO19685 Human PRO
134	41	91.1	387	6	ABO00597	AbO00597 Novel hum	207	41	91.1	1338	8	ADP54811	ADP54811 Human PRO
135	41	91.1	406	6	ABP58350	Abp58350 Human cel	208	41	91.1	1338	8	ADP24219	ADP24219 PRO polypp
136	41	91.1	413	9	ADM17172	Adm17172 Eucalyptu	209	41	91.1	1338	8	ADT07511	ADT07511 Human col
137	41	91.1	422	3	AAH41764	Aah41764 Arabidops	210	41	91.1	1340	6	AAH37017	AAH37017 Human nuc
138	41	91.1	440	4	AAU30409	Aau30409 Novel hum	211	41	91.1	1400	9	ADV85501	ADV85501 Mouse bro
139	41	91.1	465	8	ADT93239	Adt93239 Human Hun	212	41	91.1	1420	4	ABB63410	ABB63410 Drosophil
140	41	91.1	475	7	ABM87048	Abm87048 Rice abio	213	41	91.1	1430	7	ADJ69277	ADJ69277 Human hea
141	41	91.1	481	8	ADT02468	Adt02468 Human CRE	214	41	91.1	1506	9	ADZ13633	ADZ13633 Human can
142	41	91.1	485	4	ABE58326	AbE58326 Drosophil	215	41	91.1	1506	9	ADZ13622	ADZ13622 Human can
143	41	91.1	485	6	ABP58349	Abp58349 Human cel	216	41	91.1	1509	8	ADT07512	ADT07512 Human col
144	41	91.1	485	7	ADC31623	Adc31623 Human nov	217	41	91.1	1509	8	ADT07513	ADT07513 Human col
145	41	91.1	485	7	ADM05364	Adm05364 Human pro	218	41	91.1	1542	5	ABB78013	ABB78013 Amino aci
146	41	91.1	489	8	ADY10556	Ady10556 Plant ful	219	41	91.1	1616	8	ADP22958	ADP22958 PRO polypp
147	41	91.1	494	8	ADT02466	Adt02466 Human CRE	220	41	91.1	1638	7	ADC31236	ADC31236 Human nov
148	41	91.1	502	8	ADP12999	Adp12999 Protein e	221	41	91.1	1682	8	ADR14131	ADR14131 Human NF-
149	41	91.1	513	2	AAV33500	Aav33500 Human hun	222	41	91.1	1682	8	ADR14131	ADR14131 Human NF-
150	41	91.1	516	9	ADV77063	Adv77063 Huntingto	223	41	91.1	2000	8	ADP23488	ADP23488 PRO polypp
151	41	91.1	530	2	AAV33501	Aav33501 Human apo	224	41	91.1	2004	5	ABG95113	ABG95113 Human tra
152	41	91.1	535	7	ABM89001	Abm89001 Rice abio	225	41	91.1	2004	9	ADY07185	ADY07185 Cyclicin-de
153	41	91.1	549	2	AAW63672	Aaw63672 Polypteti	226	41	91.1	2004	9	ADY17185	ADY17185 PRO polypp
154	41	91.1	549	4	AAH73615	Aah73615 Human kio	227	41	91.1	2397	8	ADU02443	ADU02443 Novel hum
155	41	91.1	549	8	ADQ39810	Adq39810 Human myo	228	41	91.1	2614	9	ADZ13635	ADZ13635 Human can
156	41	91.1	552	2	AAQ33502	Aaq33502 Human apo	229	41	91.1	2614	9	ADZ13624	ADZ13624 Human can
157	41	91.1	568	3	AG411763	Ag411763 Arabidops	230	41	91.1	2910	8	ADQ59382	ADQ59382 Human can
158	41	91.1	580	8	ADT02467	Adt02467 Human CRE	231	41	91.1	2910	9	ADZ13630	ADZ13630 Human can
159	41	91.1	589	2	AAV33503	Aav33503 Human apo	232	41	91.1	2910	9	ADZ13641	ADZ13641 Human can
160	41	91.1	590	3	AG411762	Ag411762 Arabidops	233	41	91.1	2911	6	ABU56630	ABU56630 Lung canc
161	41	91.1	593	4	AAH95646	Aah95646 Human pro	234	41	91.1	2911	6	ABO07259	ABO07259 Human p53
162	41	91.1	593	6	AAH79911	Aah79911 MECT1. 4/	235	41	91.1	2911	7	ADN38844	ADN38844 Cancer/ha
163	41	91.1	597	8	ADU02790	Adu02790 Novel hum	236	41	91.1	2911	7	ADN38844	ADN38844 Human hea
164	41	91.1	629	8	ADT02461	Adt02461 mCREAP1	237	41	91.1	2911	8	ADQ18204	ADQ18204 Human sof
165	41	91.1	634	8	ABO84900	AbO84900 Human can	238	41	91.1	2911	8	ADU06413	ADU06413 Novel bro
166	41	91.1	650	4	ADU28167	Adu28167 Novel hum	239	41	91.1	2912	4	ABG06402	ABG06402 Novel hum
167	41	91.1	650	8	ADT02434	Adt02434 Human CRE	240	41	91.1	2912	9	ADZ13626	ADZ13626 Human can
168	41	91.1	657	4	AAU28355	Aau28355 Novel hum	241	41	91.1	2912	9	ADZ13637	ADZ13637 Human can
169	41	91.1	716	2	AAW38192	Aaw38192 Maize ZCA	242	41	91.1	2912	9	ADZ13628	ADZ13628 Human can
170	41	91.1	716	3	AAH28591	Aah28591 Maize ZCR	243	41	91.1	2912	9	ADZ13639	ADZ13639 Human can

244	41	91.1	3051	6	ABR58308	Abt58308	BCU0947 p	317	40	88.9	77	4	AAO2698	Human pol
245	41	91.1	3124	8	ADJ78494	Adj78494	E1A bindi	318	40	88.9	78	4	AAO2698	Human pol
246	41	91.1	3141	7	ADJ70444	Adj70444	Human hea	319	40	88.9	78	4	AAO2698	Human pol
247	41	91.1	3144	2	AAW36887	Aaw36887	Protein e	320	40	88.9	80	5	AAO2698	Human tra
248	41	91.1	3144	2	AAW36887	Aaw36887	Previous e	321	40	88.9	82	3	AAO2698	Human sec
249	41	91.1	3144	2	AAW09871	Aaw09871	Human hun	322	40	88.9	92	4	AAO2698	Human sec
250	41	91.1	3144	2	AAW44742	Aaw44742	Human hun	323	40	88.9	95	4	AAO2698	Human sec
251	41	91.1	3144	2	AAW33493	Aaw33493	Human hun	324	40	88.9	95	4	AAO2698	Human pol
252	41	91.1	3144	9	ADY98141	Ady98141	Human hun	325	40	88.9	95	5	AAO2698	DNA bindi
253	41	91.1	3223	4	ABE11407	Abel11407	Human hun	326	40	88.9	96	3	AAO2698	Zea maye
254	41	91.1	3223	4	ABE11470	Abel11470	Human hun	327	40	88.9	97	8	AAO2698	Human EST
255	41	91.1	3469	8	ADG96634	Adg96634	Drosophil	328	40	88.9	98	4	AAO2698	Novel hum
256	41	91.1	3502	4	ABE58382	Abel58382	Drosophil	329	40	88.9	98	5	AAO2698	Human bre
257	41	91.1	3604	8	ADO89594	Adg89594	Antagonis	330	40	88.9	98	5	AAO2698	Human sec
258	40	88.9	9	6	ABR28352	Abt28352	Human can	331	40	88.9	99	5	ABP08829	Human ORF
259	40	88.9	10	6	ABR28347	Abt28347	Human can	332	40	88.9	100	4	ABG25413	Novel hum
260	40	88.9	15	2	AAW39016	Aaw39016	Peptide r	333	40	88.9	106	4	ABG07829	Novel hum
261	40	88.9	15	2	AAW38985	Aaw38985	Peptide r	334	40	88.9	107	4	AAO00808	Human pol
262	40	88.9	15	6	ABR38294	Abt38294	Human can	335	40	88.9	107	6	ADA55418	Human pro
263	40	88.9	15	6	ABR38391	Abt38391	Human can	336	40	88.9	108	4	ABG69539	Drosophil
264	40	88.9	15	6	ABR38393	Abt38393	Human can	337	40	88.9	108	5	ABG62148	Human pro
265	40	88.9	15	6	ABR38392	Abt38392	Human can	338	40	88.9	109	4	AAO01024	Human pol
266	40	88.9	15	6	ABR38351	Abt38351	Human can	339	40	88.9	110	4	ABG68326	Drosophil
267	40	88.9	15	6	ABR38292	Abt38292	Human can	340	40	88.9	110	4	AAO05822	Human pro
268	40	88.9	30	4	AAO02051	Aao02051	Human pol	341	40	88.9	112	4	AAO02189	Human pol
269	40	88.9	33	4	AAO12108	Aao12108	Human pol	342	40	88.9	113	5	ABP57768	Human zin
270	40	88.9	37	6	AAO19948	Aao19948	Alzheimer	343	40	88.9	114	4	AAO05912	Human pol
271	40	88.9	38	6	AAO19947	Aao19947	Alzheimer	344	40	88.9	114	4	AAO02218	Human pol
272	40	88.9	41	1	AAO10196	Aao10196	Sequence	345	40	88.9	114	7	ADC33200	Human nov
273	40	88.9	41	1	AAO05676	Aao05676	Human pol	346	40	88.9	115	4	AAU32525	Novel hum
274	40	88.9	42	1	AAO71006	Aap71006	Sequence	347	40	88.9	115	4	AAU32143	Novel hum
275	40	88.9	42	1	AAO80040	Aap80040	Beta-huma	348	40	88.9	115	5	ABR01725	Human bre
276	40	88.9	42	1	AAO91846	Aap91846	Analogue	349	40	88.9	116	4	ABG66710	Drosophil
277	40	88.9	42	1	AAW42213	Aaw42213	Peptide d	350	40	88.9	116	4	AAO00466	Human pol
278	40	88.9	42	2	AAW69459	Aaw69459	HCG antiq	351	40	88.9	116	5	ADG79555	Human sec
279	40	88.9	42	2	AAW93444	Aay93444	Human hCG	352	40	88.9	116	5	ADG79456	Human sec
280	40	88.9	42	3	AAW87489	Aay87489	Human cho	353	40	88.9	116	8	ADG66101	Novel hum
281	40	88.9	42	3	AAW20553	Aao20553	Human cho	354	40	88.9	117	7	ADD19292	Human sec
282	40	88.9	42	4	AAU01149	Aau01149	Structure	355	40	88.9	117	7	ADD19292	Human sec
283	40	88.9	42	4	AAW48395	Aaw48395	Human cho	356	40	88.9	118	7	ADB37548	Neural th
284	40	88.9	42	4	AAU02847	Aau02847	Human cho	357	40	88.9	120	8	ADX70886	Plant ful
285	40	88.9	42	4	AAW04131	Aab04131	Peptide f	358	40	88.9	121	4	AAO13020	Human pol
286	40	88.9	42	5	AAW50788	Aam50788	Human cho	359	40	88.9	121	4	ABG16589	Novel hum
287	40	88.9	46	4	AAO10556	Aao10556	Human pol	360	40	88.9	125	4	AAO04869	Human pol
288	40	88.9	46	4	ABG17583	Abg17583	Novel hum	361	40	88.9	125	7	ADD19266	Human sec
289	40	88.9	48	1	AAO10195	Abp10195	Sequence	362	40	88.9	126	4	AAO00766	Human pol
290	40	88.9	48	1	AAO80039	Aap80039	Beta-huma	363	40	88.9	130	8	ADR08968	Human pro
291	40	88.9	48	1	AAO91845	Aap91845	Analogue	364	40	88.9	131	4	AAO01747	Human pol
292	40	88.9	48	2	AAW42212	Aaw42212	Peptide d	365	40	88.9	132	4	AAO00570	Human pol
293	40	88.9	48	2	AAW69458	Aaw69458	HCG antiq	366	40	88.9	132	4	AAO03924	Human pol
294	40	88.9	48	2	AAW93443	Aay93443	Human hCG	367	40	88.9	135	7	ADB08953	Novel pro
295	40	88.9	48	3	AAW87488	Aay87488	Human cho	368	40	88.9	135	8	ADS11754	Human the
296	40	88.9	48	3	AAW20552	Aab20552	Human cho	369	40	88.9	136	4	AAO01204	Human pol
297	40	88.9	48	4	AAU01148	Aau01148	Structure	370	40	88.9	137	4	ABG13622	Novel hum
298	40	88.9	48	4	AAW48394	Aaw48394	Human cho	371	40	88.9	138	8	ADG66021	Novel hum
299	40	88.9	48	4	AAU02846	Aau02846	Human cho	372	40	88.9	141	4	AAO05999	Human pol
300	40	88.9	48	4	AAW04130	Aab04130	Peptide f	373	40	88.9	144	5	ADK37055	Novel hum
301	40	88.9	48	5	ADK36186	Adk36186	Novel hum	374	40	88.9	144	5	ADK37004	Novel hum
302	40	88.9	48	5	ADK35869	Adk35869	Novel hum	375	40	88.9	146	6	ABP75894	Human sec
303	40	88.9	50	4	AAU49050	Aau49050	Propionib	376	40	88.9	148	8	ADG65118	Novel hum
304	40	88.9	50	6	ABW45569	Abm45569	Propionib	377	40	88.9	148	8	ADG66558	Novel hum
305	40	88.9	52	4	AAO11780	Aao11780	Human pol	378	40	88.9	149	4	AAO03286	Human pol
306	40	88.9	53	2	AAO26916	Aay26916	MLV MCF e	379	40	88.9	152	4	ABG16591	Novel hum
307	40	88.9	56	4	AAO05154	Aao05154	Human pol	380	40	88.9	152	8	ADG66482	Novel hum
308	40	88.9	56	4	ABG12672	Abg12672	Novel hum	381	40	88.9	154	4	ABG08425	Novel hum
309	40	88.9	57	4	AAO04578	Aao04578	Human pol	382	40	88.9	155	5	AAU72798	Human ant
310	40	88.9	59	4	AAO05645	Aao05645	Human pol	383	40	88.9	156	4	ABG10943	Novel hum
311	40	88.9	63	4	AAO10207	Aao10207	Human pol	384	40	88.9	158	6	ADG83852	Human POM
312	40	88.9	64	5	ABP05328	Abp05328	Human ORF	385	40	88.9	164	4	AAU27698	Full-leng
313	40	88.9	73	7	ABM87180	Abm87180	Rice abio	386	40	88.9	172	4	ABG19542	Novel hum
314	40	88.9	75	4	ABG69169	Aab69169	Rat VGF p	387	40	88.9	174	3	AAW43208	Human ORF
315	40	88.9	75	4	ABG69547	Aab69547	Rat VGFp	388	40	88.9	175	4	AAU21256	Human nov
316	40	88.9	77	4	AAO01134	Aao01134	Human pol	389	40	88.9	180	7	ADB64804	Human pro

390	40	88.9	184	8	ADI42026	Adi42026 Plant tra	463	40	88.9	265	7	ADB85017	Human PRO
391	40	88.9	184	8	ADO02576	Ado02576 Thalecres	464	40	88.9	265	7	ADB78123	Novel hum
392	40	88.9	184	8	ADO62215	Ado62215 Transcrip	465	40	88.9	265	7	ADB87189	Human PRO
393	40	88.9	197	4	ABG22105	Abg22105 Novel hum	466	40	88.9	265	7	ADB84771	Human PRO
394	40	88.9	201	4	ABG19534	Abg19534 Novel hum	467	40	88.9	265	7	ADB83886	Novel hum
395	40	88.9	207	4	AU27870	Au27870 Contig po	468	40	88.9	265	7	ADB73041	Novel hum
396	40	88.9	217	8	ADI42771	Adi42771 Plant tra	469	40	88.9	265	7	ADC57703	Human PRO
397	40	88.9	218	7	ADC86939	Adc86939 Human GPC	470	40	88.9	265	7	ADC55067	Human PRO
398	40	88.9	219	4	AG666500	Ag666500 Human neu	471	40	88.9	265	7	ADC11934	Human sec
399	40	88.9	220	7	ABM85689	Abm85689 Human pro	472	40	88.9	265	7	ADC56356	Human PRO
400	40	88.9	220	4	AAU30601	Aau30601 Novel hum	473	40	88.9	265	7	ADC07411	Human sec
401	40	88.9	228	4	AAB60487	Aab60487 Human cel	474	40	88.9	265	7	ADC11401	Human sec
402	40	88.9	230	5	ABP43974	Abp43974 FLJ23018	475	40	88.9	265	7	ADC36879	Human PRO
403	40	88.9	232	4	AM24472	Am24472 Human BST	476	40	88.9	265	7	ADC21869	Human PRO
404	40	88.9	232	7	ADI60378	Adi60378 Secreted	477	40	88.9	265	7	ADC49900	Novel hum
405	40	88.9	234	4	ABG24732	Abg24732 Novel hum	478	40	88.9	265	7	ADC49099	Novel hum
406	40	88.9	235	6	ABJ19682	Abj19682 Human sec	479	40	88.9	265	7	ADC49616	Novel hum
407	40	88.9	235	6	ABP99572	Abp99572 Human sec	480	40	88.9	265	7	ADC49616	Novel hum
408	40	88.9	235	8	ADT59093	Adt59093 Plant pol	481	40	88.9	265	7	ADC47477	Novel hum
409	40	88.9	236	3	ABX39216	Abx39216 Human sec	482	40	88.9	265	7	ADC14523	Novel hum
410	40	88.9	238	7	ABM89261	Abm89261 Rice abio	483	40	88.9	265	7	ADC47222	Novel hum
411	40	88.9	245	8	ADT59408	Adt59408 Plant pol	484	40	88.9	265	7	ADD08055	Novel hum
412	40	88.9	247	4	ABG15898	Abg15898 Novel hum	485	40	88.9	265	7	ADC81880	Human PRO
413	40	88.9	247	4	ADG89176	Adg89176 Human sec	486	40	88.9	265	7	ADD07522	Novel hum
414	40	88.9	252	8	ADX96980	Adx96980 Plant ful	487	40	88.9	265	7	ADC78097	Novel hum
415	40	88.9	261	7	ABM89090	Abm89090 Rice abio	488	40	88.9	265	7	ADC82413	Human PRO
416	40	88.9	262	8	ADY08954	Ady08954 Plant ful	489	40	88.9	265	7	ADD06332	Novel hum
417	40	88.9	265	3	AA766691	Aay66691 Membrane-	490	40	88.9	265	7	ADD08593	Novel hum
418	40	88.9	265	3	AAB24063	Aab24063 Human PRO	491	40	88.9	265	7	ADC77851	Novel hum
419	40	88.9	265	4	AAB65214	Aab65214 Human PRO	492	40	88.9	265	7	ADD06842	Novel hum
420	40	88.9	265	5	AU83666	Au83666 Human PRO	493	40	88.9	265	7	ADC83089	Human PRO
421	40	88.9	265	5	ADY31890	Ady31890 Novel hum	494	40	88.9	265	7	ADD50814	Novel hum
422	40	88.9	265	6	ABU58029	Abu58029 Human PRO	495	40	88.9	265	7	ADD51060	Novel hum
423	40	88.9	265	6	ABU59107	Abu59107 Novel hum	496	40	88.9	265	7	ADD55196	Human PRO
424	40	88.9	265	6	ABU82619	Abu82619 Human sec	497	40	88.9	265	7	ADD50541	Human PRO
425	40	88.9	265	6	ABU60538	Abu60538 Human sec	498	40	88.9	265	7	ADD54592	Human PRO
426	40	88.9	265	6	ABU80813	Abu80813 Human PRO	499	40	88.9	265	7	ADD50295	Human PRO
427	40	88.9	265	6	ABO33779	Ab033779 Novel hum	500	40	88.9	265	7	ADD51306	Novel hum
428	40	88.9	265	6	ABU13920	Abu13920 Human PRO	501	40	88.9	265	7	ADD51306	Novel hum
429	40	88.9	265	6	ABU72505	Abu72505 Novel hum	502	40	88.9	265	7	ADD51306	Novel hum
430	40	88.9	265	6	ABU59254	Abu59254 Human sec	503	40	88.9	265	7	ADD51306	Novel hum
431	40	88.9	265	6	ABO25951	Ab025951 Human PRO	504	40	88.9	265	7	ADD51306	Novel hum
432	40	88.9	265	6	ABU82122	Abu82122 Novel hum	505	40	88.9	265	7	ADD51306	Novel hum
433	40	88.9	265	6	ABU58960	Abu58960 Human sec	506	40	88.9	265	7	ADD51306	Novel hum
434	40	88.9	265	6	ABU92338	Abu92338 Novel hum	507	40	88.9	265	7	ADD51306	Novel hum
435	40	88.9	265	6	ABU59403	Abu59403 Novel hum	508	40	88.9	265	7	ADD51306	Novel hum
436	40	88.9	265	6	ABU92169	Abu92169 Novel hum	509	40	88.9	265	7	ADD51306	Novel hum
437	40	88.9	265	6	ABU10875	Abu10875 Human PRO	510	40	88.9	265	7	ADD51306	Novel hum
438	40	88.9	265	6	ABU81627	Abu81627 Novel hum	511	40	88.9	265	7	ADD51306	Novel hum
439	40	88.9	265	6	ABU88566	Abu88566 Human sec	512	40	88.9	265	7	ADD51306	Novel hum
440	40	88.9	265	6	ABO34080	Ab034080 Human PRO	513	40	88.9	265	7	ADD51306	Novel hum
441	40	88.9	265	6	ABJ72302	Abj72302 Human PRO	514	40	88.9	265	7	ADD51306	Novel hum
442	40	88.9	265	6	ADA37734	Ada37734 Human sec	515	40	88.9	265	7	ADD51306	Novel hum
443	40	88.9	265	6	ADA21420	Ada21420 Human sec	516	40	88.9	265	7	ADD51306	Novel hum
444	40	88.9	265	6	ADA10207	Ada10207 Human sec	517	40	88.9	265	7	ADD51306	Novel hum
445	40	88.9	265	6	ADA17751	Ada17751 Human PRO	518	40	88.9	265	7	ADD51306	Novel hum
446	40	88.9	265	6	ADA27859	Ada27859 Human sec	519	40	88.9	265	7	ADD51306	Novel hum
447	40	88.9	265	6	ADA94439	Ada94439 Human sec	520	40	88.9	265	7	ADD51306	Novel hum
448	40	88.9	265	6	ADA38664	Ada38664 Human sec	521	40	88.9	265	7	ADD51306	Novel hum
449	40	88.9	265	6	ABJ72430	Abj72430 Human PRO	522	40	88.9	265	7	ADD51306	Novel hum
450	40	88.9	265	6	ABJ92785	Abj92785 Human sec	523	40	88.9	265	7	ADD51306	Novel hum
451	40	88.9	265	6	ABO34325	Ab034325 Human sec	524	40	88.9	265	7	ADD51306	Novel hum
452	40	88.9	265	7	ABO53166	Ab053166 Human sec	525	40	88.9	265	7	ADD51306	Novel hum
453	40	88.9	265	7	ADA22346	Ada22346 Human sec	526	40	88.9	265	7	ADD51306	Novel hum
454	40	88.9	265	7	ADA02536	Ada02536 Human sec	527	40	88.9	265	7	ADD51306	Novel hum
455	40	88.9	265	7	ADA06512	Ada06512 Human sec	528	40	88.9	265	7	ADD51306	Novel hum
456	40	88.9	265	7	ABJ72132	Abj72132 Human mem	529	40	88.9	265	7	ADD51306	Novel hum
457	40	88.9	265	7	ADA39205	Ada39205 Human sec	530	40	88.9	265	7	ADD51306	Novel hum
458	40	88.9	265	7	ADB83640	Abd83640 Novel hum	531	40	88.9	265	7	ADD51306	Novel hum
459	40	88.9	265	7	ADB80746	Abd80746 Novel hum	532	40	88.9	265	7	ADD51306	Novel hum
460	40	88.9	265	7	ADB73287	Abd73287 Novel hum	533	40	88.9	265	7	ADD51306	Novel hum
461	40	88.9	265	7	ADB96231	Abd96231 Human PRO	534	40	88.9	265	7	ADD51306	Novel hum
462	40	88.9	265	7	ADB78369	Abd78369 Novel hum	535	40	88.9	265	7	ADD51306	Novel hum

536	40	88.9	265	8	ADD77133	Novel hum	609	40	88.9	359	9	ADV68459	Adv68459	NHK amino
537	40	88.9	265	8	ADD85827	Novel hum	610	40	88.9	359	9	ADY54996	Ady54996	Human nov
538	40	88.9	265	8	AD505376	Human PRO	611	40	88.9	361	7	AM590394	Am590394	Rice abio
539	40	88.9	265	8	ADD74851	Human PRO	612	40	88.9	362	7	ABM87638	Abm87638	Rice abio
540	40	88.9	265	8	ADDF35349	Human PRO	613	40	88.9	363	4	AM25490	Aam25490	Human pro
541	40	88.9	265	8	ADG115599	Human PRO	614	40	88.9	371	7	ABO43147	Abo43147	A. thalia
542	40	88.9	265	8	ADG05663	Novel hum	615	40	88.9	371	8	ADO02519	Ado02519	Thalecres
543	40	88.9	265	8	ADG27217	Human PRO	616	40	88.9	373	8	AAU03507	Aau03507	Human pro
544	40	88.9	265	8	ADG11280	Novel hum	617	40	88.9	373	4	ABE64723	AbE64723	Drosophil
545	40	88.9	265	8	ADG12059	Novel hum	618	40	88.9	389	6	ABU33914	Abu33914	Protein e
546	40	88.9	265	8	ADF94616	Novel hum	619	40	88.9	390	2	AA17066	Aay17066	Human 3-O
547	40	88.9	265	8	ADG06712	Human PRO	620	40	88.9	390	4	AA95507	Aa95507	Human pro
548	40	88.9	265	8	ADH39056	Novel hum	621	40	88.9	390	8	ADI09937	Adi09937	Human hep
549	40	88.9	265	8	ADH19469	Human sec	622	40	88.9	390	8	ADJ76166	Adj76166	Marker ge
550	40	88.9	265	8	ADH20962	Human sec	623	40	88.9	390	8	ADL61251	Adl61251	Human hep
551	40	88.9	265	8	ADH20002	Human sec	624	40	88.9	390	9	ADX26316	Adx26316	Novel cel
552	40	88.9	265	8	ADG34146	Novel hum	625	40	88.9	391	4	AA70114	Aab70114	Murine 3-
553	40	88.9	265	8	ADG33616	Human PRO	626	40	88.9	391	4	AA70115	Aab70115	Human 3-O
554	40	88.9	265	8	ADH69710	Human PRO	627	40	88.9	399	3	AA808218	Aab08218	A protein
555	40	88.9	265	8	ADI29871	Novel hum	628	40	88.9	402	9	ADX83401	Adx83401	Human TSG
556	40	88.9	265	8	ADM27268	Novel hum	629	40	88.9	404	8	ADT58544	Adt58544	Plant pol
557	40	88.9	265	8	ADK66626	Human PRO	630	40	88.9	412	6	ABR42686	AbR42686	Petunia a
558	40	88.9	265	9	AEA38486	Human sec	631	40	88.9	414	6	AAE32027	Aae32027	Human kin
559	40	88.9	266	5	AA48706	Human tra	632	40	88.9	414	8	ADJ96641	Adj96641	Human Nim
560	40	88.9	267	8	ADY11763	Plant ful	633	40	88.9	416	5	AAE15660	Aae15660	Human hom
561	40	88.9	268	8	AD002317	Thalecres	634	40	88.9	416	6	ABP97648	Abp97648	Amino aci
562	40	88.9	270	8	ADY09279	Plant ful	635	40	88.9	416	9	ABE48139	Abe48139	Amino aci
563	40	88.9	275	8	ADQ66448	Novel hum	636	40	88.9	431	3	AA32324	Aag32324	Arabidops
564	40	88.9	285	7	AD160139	Secreted	637	40	88.9	431	3	AA309613	Aag09613	Arabidops
565	40	88.9	291	7	ABM88444	Rice abio	638	40	88.9	431	8	ADI43535	Adi43535	Plant tra
566	40	88.9	309	7	ADC31415	Human nov	639	40	88.9	431	9	AEA26765	Aea26765	Stress to
567	40	88.9	310	7	ABM74160	DNA clone	640	40	88.9	432	5	ABP69406	Abp69406	Human pol
568	40	88.9	311	4	AAW24509	C883P pre	641	40	88.9	432	5	ABU65180	Abu65180	Human NOV
569	40	88.9	311	5	AA47945	CDX2. 6/2	642	40	88.9	432	8	ADN62011	Adn62011	Human nov
570	40	88.9	313	6	ABP55358	Human col	643	40	88.9	432	9	ADY18827	Ady18827	PRO polyP
571	40	88.9	313	8	AD089848	Antagonis	644	40	88.9	434	9	ABM91399	Abm91399	M. xanthu
572	40	88.9	316	5	AAE22089	Human nov	645	40	88.9	435	3	AA309612	Aag09612	Arabidops
573	40	88.9	320	3	AA346813	Arabidops	646	40	88.9	439	7	ADP74169	Adp74169	Human nov
574	40	88.9	320	8	ADG37027	Arabidops	647	40	88.9	439	8	ADY25088	Ady25088	Plant ful
575	40	88.9	320	8	ADR09376	Human PRO	648	40	88.9	441	3	AA349463	Aag49463	Arabidops
576	40	88.9	322	7	ABM86610	Rice abio	649	40	88.9	441	6	ABP75500	Abp75500	Human sec
577	40	88.9	323	8	ADI43438	Plant tra	650	40	88.9	443	3	AA349462	Aag49462	Arabidops
578	40	88.9	329	8	ADX80067	Plant ful	651	40	88.9	453	7	ADD48354	Ad448354	Human PRO
579	40	88.9	332	7	ABM88621	Rice abio	652	40	88.9	455	7	ABO69519	AbO69519	Pseudomon
580	40	88.9	333	8	AD089903	Human PRO	653	40	88.9	456	7	ADE40443	AdE40443	Human cho
581	40	88.9	334	5	ABJ04705	Mycobacte	654	40	88.9	461	4	AAU17200	Aau17200	Novel sig
582	40	88.9	335	7	ABM86681	Rice abio	655	40	88.9	461	7	ADB93908	AdB93908	Human nov
583	40	88.9	339	7	ABM86928	Rice abio	656	40	88.9	461	7	ADJ70999	Adj70999	Human hea
584	40	88.9	343	3	AA841666	Human ORF	657	40	88.9	461	8	ADY08284	Ady08284	Plant ful
585	40	88.9	345	5	AAU84271	Human end	658	40	88.9	463	3	AA41079	Aa41079	Human ORF
586	40	88.9	346	2	AAW14442	Human tra	659	40	88.9	471	8	ADJ71959	Adj71959	Human PMW
587	40	88.9	346	7	ADDA6201	Human PRO	660	40	88.9	477	6	ABP72344	Abp72344	Brain fac
588	40	88.9	346	7	ADBE2409	Human PRO	661	40	88.9	477	7	ADJ80175	Adj80175	Novel hum
589	40	88.9	346	9	ADX05441	Cyclin-de	662	40	88.9	480	2	AA44551	Aar44551	Brain fac
590	40	88.9	350	7	ADG42159	Human bra	663	40	88.9	480	4	ABG12623	Abg12623	Novel hum
591	40	88.9	350	7	ADF74180	Human nov	664	40	88.9	481	5	ABE57076	AbE57076	Mouse isc
592	40	88.9	350	9	AEA50208	Novel hum	665	40	88.9	483	8	ADU02754	AdU02754	Novel hum
593	40	88.9	352	7	ADM04873	Human PRO	666	40	88.9	485	7	ADD30204	AdD30204	Plant yie
594	40	88.9	353	5	AAE22090	Human nov	667	40	88.9	485	8	ADI41561	Adi41561	Plant tra
595	40	88.9	354	4	AA860375	Human ato	668	40	88.9	485	8	AEA26633	Aea26633	Stress to
596	40	88.9	354	4	AA860349	Human ato	669	40	88.9	486	8	ADN72955	Adn72955	Thale cre
597	40	88.9	354	6	AAO30945	Human ato	670	40	88.9	489	8	ADQ19922	AdQ19922	Human bof
598	40	88.9	354	6	AAO30920	Human ato	671	40	88.9	491	8	ADY07141	Ady07141	Plant ful
599	40	88.9	354	8	ADQ19897	Human sof	672	40	88.9	492	8	ADE28313	Ade28313	Human KPP
600	40	88.9	354	8	ADR69062	Human ato	673	40	88.9	492	8	ADR86074	Adr86074	Aspergill
601	40	88.9	354	8	ADR69063	Human ato	674	40	88.9	492	8	ADU02730	AdU02730	Novel hum
602	40	88.9	356	4	AA870701	Human TAA	675	40	88.9	493	4	ABG12673	Abg12673	Novel hum
603	40	88.9	357	5	AAU85555	Clone #50	676	40	88.9	494	3	AA69386	Aay69386	Amino aci
604	40	88.9	357	5	AAU69431	Lung smal	677	40	88.9	494	3	AA69393	Aay69393	Amino aci
605	40	88.9	357	6	ABU69527	Human lun	678	40	88.9	498	8	ADX87354	Adx87354	Plant ful
606	40	88.9	357	6	ABU66430	Lung canc	679	40	88.9	500	4	ABG03058	Abg03058	Novel hum
607	40	88.9	357	7	ADH47336	Human lun	680	40	88.9	500	4	ABG19840	Abg19840	Novel hum
608	40	88.9	357	8	ADJ21255	Human lun	681	40	88.9	500	6	ABR01808	AbR01808	Human can

682	40	88.9	503	7	ABR62851	Abr62851 Murine To	755	40	88.9	690	7	ADM29409	Human nov
683	40	88.9	505	8	ADT59741	Adt59741 Plant pol	756	40	88.9	690	7	ADM29397	Human nov
684	40	88.9	519	8	ADW17238	Adw17238 Eucalyptu	757	40	88.9	690	7	ADM29389	Human nov
685	40	88.9	522	8	ADW17370	Adw17370 Plant ful	758	40	88.9	695	7	ADM29399	Human nov
686	40	88.9	522	8	ADX93027	Adx93027 Plant ful	759	40	88.9	696	7	ADM29407	Human nov
687	40	88.9	523	7	ADC87177	Adc87177 Human Gpc	760	40	88.9	706	7	ADM05234	Human pro
688	40	88.9	524	3	RAG17154	Rag17154 Arabidops	761	40	88.9	708	7	ADC19791	Human can
689	40	88.9	524	7	ADC31173	Adc31173 Human nov	762	40	88.9	714	8	ADY06938	Plant ful
690	40	88.9	523	8	ADX96523	Adx96523 Plant ful	763	40	88.9	717	8	ADC01844	C. albica
691	40	88.9	535	8	ADX91150	Adx91150 Plant ful	764	40	88.9	721	5	AAE22152	Human TRN
692	40	88.9	543	6	AAE30109	Aae30109 Heliothis	765	40	88.9	721	5	ADM04785	Human pro
693	40	88.9	546	2	AAW39788	Aaw39788 Tobacco P	766	40	88.9	721	7	ADM29395	Human nov
694	40	88.9	547	7	ABM89005	Abm89005 Rice abio	767	40	88.9	737	9	ADV99974	Human PEM
695	40	88.9	549	8	ADX88993	Adx88993 Plant ful	768	40	88.9	737	9	ADV99972	Human PEM
696	40	88.9	553	4	ABB60817	Abb60817 Drosophil	769	40	88.9	738	4	ADM29393	Human nov
697	40	88.9	553	9	ABM90608	Abm90608 M. xanthu	770	40	88.9	748	4	AAAM40227	Human pol
698	40	88.9	557	8	ADL57099	Adl57099 Human NOV	771	40	88.9	771	4	ABB64331	Drosophil
699	40	88.9	562	7	ADB65060	Adb65060 Human pro	772	40	88.9	775	5	AAU93161	Arabidops
700	40	88.9	562	8	ADR58973	Adr58973 Human Elk	773	40	88.9	780	6	AAE33674	Human str
701	40	88.9	562	8	ADR58971	Adr58971 Human Elk	774	40	88.9	781	7	ADM29391	Human nov
702	40	88.9	570	7	ADB64564	Adb64564 Human pro	775	40	88.9	804	4	ABG03827	Novel hum
703	40	88.9	581	5	ABG66745	Abg66745 Human nov	776	40	88.9	804	7	ADC33210	Human nov
704	40	88.9	585	4	AAW93788	Aaw93788 Human pol	777	40	88.9	814	8	ADK70212	Human oes
705	40	88.9	585	8	ADL31778	Adl31778 Human pro	778	40	88.9	814	8	ADL72166	Human sol
706	40	88.9	585	8	ABM81730	Abm81730 Tumour-as	779	40	88.9	817	5	AAO15419	Novel hum
707	40	88.9	589	5	ABP73306	Abp73306 Candida a	780	40	88.9	828	4	ABBS6368	Drosophil
708	40	88.9	594	4	ABB61362	Abb61362 Drosophil	781	40	88.9	834	5	ABP74067	Candida a
709	40	88.9	600	8	ADO62366	Ado62366 Transcript	782	40	88.9	863	8	ADI43159	Plant tra
710	40	88.9	603	5	AAE22088	Aae22088 Human nov	783	40	88.9	863	8	ADO03242	Thalecres
711	40	88.9	616	3	ADJ65163	Adj65163 Potassium	784	40	88.9	863	8	ADO62535	Transcrip
712	40	88.9	603	8	ADT58720	Adt58720 Plant pol	785	40	88.9	876	8	ABO84874	Human can
713	40	88.9	615	8	ADQ19437	Adq19437 Human sof	786	40	88.9	885	2	AAR99740	Human can
714	40	88.9	615	8	ABM80918	Abm80918 Tumour-as	787	40	88.9	914	4	ABBB66394	Drosophil
715	40	88.9	616	9	AAE17692	Aae17692 Alzheimer	788	40	88.9	933	9	AEA00186	Human TAT
716	40	88.9	616	4	AAU09069	Aau09069 Human neu	789	40	88.9	933	9	AEA00706	Human TAT
717	40	88.9	616	6	AAO19950	Aao19950 Alzheimer	790	40	88.9	965	7	ADC31607	Human nov
718	40	88.9	616	7	ADD14195	Add14195 Human src	791	40	88.9	978	7	ADC31742	Human nov
719	40	88.9	616	9	ADY70363	Ady70363 Human bet	792	40	88.9	1026	3	AAI83025	Staufen p
720	40	88.9	616	9	AAE17683	Aae17683 Alzheimer	793	40	88.9	1026	4	ABB61775	Drosophil
721	40	88.9	617	7	ADE83413	Ade83413 Rat Prote	794	40	88.9	1045	4	ABB62505	Drosophil
722	40	88.9	617	7	ADE58004	Ade58004 Rat Prote	795	40	88.9	1051	2	AAI29321	Human ata
723	40	88.9	622	4	ABB62816	Abb62816 Drosophil	796	40	88.9	1052	6	AAE33773	Human nuc
724	40	88.9	622	5	ABG79656	Abg79656 Invertebr	797	40	88.9	1052	7	ADG71258	Novel hum
725	40	88.9	622	8	ADJ65161	Adj65161 Potassium	798	40	88.9	1059	8	ADN99406	Novel hum
726	40	88.9	622	9	ABE44574	Abe44574 Sleep dis	799	40	88.9	1060	8	ADS10464	Human the
727	40	88.9	639	8	ADF77438	Adf77438 MLV Env p	800	40	88.9	1063	8	ADR09556	Human pro
728	40	88.9	641	8	ADF77489	Adf77489 MLV Env p	801	40	88.9	1070	8	ADX05832	Cyclin-de
729	40	88.9	642	4	AAE08927	Aae08927 Human NOV	802	40	88.9	1105	8	ADT04693	Rice DNA
730	40	88.9	642	7	ADM29403	Adm29403 Human nov	803	40	88.9	1117	8	ADI82548	Human mod
731	40	88.9	645	2	AAI28817	Aai28817 pt326_4 s	804	40	88.9	1134	5	AAE24341	Human lun
732	40	88.9	645	4	AAW78331	Aaw78331 Human pro	805	40	88.9	1134	7	ADF74123	Human nov
733	40	88.9	645	4	AAW39207	Aam39207 Human pol	806	40	88.9	1232	4	ABB58856	Drosophil
734	40	88.9	645	4	AAW94070	Aab94070 Human pro	807	40	88.9	1242	6	ABG72372	Human Ins
735	40	88.9	650	2	AAE71382	Aae71382 Drosophil	808	40	88.9	1242	2	AAW93972	Human Ins
736	40	88.9	650	8	ADP04179	Adp04179 Human col	809	40	88.9	1242	4	AAE83921	Amino aci
737	40	88.9	655	3	AAI96448	Aai96448 Forkhead	810	40	88.9	1242	6	ABG99797	Amino aci
738	40	88.9	655	3	AAE06076	Aae06076 Human hom	811	40	88.9	1242	6	ABG72369	Human Ins
739	40	88.9	655	8	ADN06032	Adn06032 Antipsoric	812	40	88.9	1242	6	ABG72372	Human Ins
740	40	88.9	655	8	ADN022517	Adn022517 Biochemic	813	40	88.9	1242	6	ABG72370	Human Ins
741	40	88.9	655	8	ADN40823	Adn40823 Human for	814	40	88.9	1242	6	ABG72371	Human Ins
742	40	88.9	655	8	ABM81512	Abm81512 Tumour-as	815	40	88.9	1242	6	ABG72368	Human Ins
743	40	88.9	655	9	ADY14882	Ady14882 PRO polyyp	816	40	88.9	1242	6	ADA00642	Human IRS
744	40	88.9	656	7	ADD47623	Add47623 Human Pro	817	40	88.9	1242	8	ADJ66659	IRSI prot
745	40	88.9	656	7	ADD48792	Add48792 Human Pro	818	40	88.9	1242	8	ADN98329	Human Ins
746	40	88.9	656	9	ADV99949	Adv99949 Human PEM	819	40	88.9	1242	8	ADO44347	Human IRS
747	40	88.9	659	9	ADV99948	Adv99948 Human PEM	820	40	88.9	1242	8	ADU87427	Human IRS
748	40	88.9	659	9	ADV99945	Adv99945 Human PEM	821	40	88.9	1242	9	ADU87427	Human IRS
749	40	88.9	669	2	AAR86408	Aar86408 Human mat	822	40	88.9	1242	9	ADX07852	Cyclin-de
750	40	88.9	669	4	AAE84617	Aae84617 Amino aci	823	40	88.9	1242	9	ADX86115	Human ins
751	40	88.9	669	4	AAE10424	Aae10424 Human mat	824	40	88.9	1242	9	ADZ79340	Human ins
752	40	88.9	669	4	AAO12939	Aao12939 Human pol	825	40	88.9	1242	9	AAE81147	Human ins
753	40	88.9	690	5	AAE22154	Aae22154 Human TRN	826	40	88.9	1243	2	AAR67708	Insulin r
754	40	88.9	690	6	ADA54663	Ada54663 Human pro	827	40	88.9	1297	7	ABM90342	Rice abio

828	40	88.9	1316	4	ABG22997	Abg22997	Novel hum	901	39	86.7	104	4	ABB18083	Abb18083	Protein #
829	40	88.9	1316	7	ADK40903	Adk40903	Novel hum	902	39	86.7	104	4	AAM65795	Aam65795	Human bon
830	40	88.9	1316	8	ADR15627	Adr15627	Kinase 36	903	39	86.7	104	4	AAM53415	Aam53415	Human bra
831	40	88.9	1316	9	AEB32334	Aeb32334	Human pro	904	39	86.7	104	4	ABG47440	Abg47440	Human liv
832	40	88.9	1316	9	AEB32358	Aeb32358	Human pro	905	39	86.7	104	4	AAM01407	Aam01407	Peptide #
833	40	88.9	1321	4	ABG25416	Abg25416	Novel hum	906	39	86.7	104	5	ABG35427	Abg35427	Human pep
834	40	88.9	1437	7	ADJ69580	Adj69580	Human hea	907	39	86.7	107	3	ABG35427	Abg35427	Fragment
835	40	88.9	1437	7	ADJ69918	Adj69918	Human hea	908	39	86.7	107	4	AAO12356	Aao12356	Human pol
836	40	88.9	1439	7	ABM88861	Abm88861	Rice abio	909	39	86.7	109	4	AAO02138	Aao02138	Human pol
837	40	88.9	1640	7	ADC59314	Adc59314	Human pol	910	39	86.7	110	4	AAO01457	Aao01457	Human pol
838	40	88.9	1703	4	ABB66223	Abb66223	Drosophil	911	39	86.7	110	8	ABO56225	Abo56225	Human gen
839	40	88.9	1798	3	AAV51611	Aav51611	Human HSG	912	39	86.7	119	4	AAO07689	Aao07689	Human pol
840	40	88.9	1819	7	ABM85839	Abm85839	Human pro	913	39	86.7	126	3	AAG56294	Aag56294	Arabidops
841	40	88.9	1963	4	ABM62819	Abm62819	Drosophil	914	39	86.7	126	8	ADT56036	Adt56036	Plant pol
842	40	88.9	1991	4	ABM60651	Abm60651	Drosophil	915	39	86.7	127	5	AAE29154	Aae29154	Human non
843	40	88.9	2075	4	ABBE1677	Abbe1677	Drosophil	916	39	86.7	128	8	ADT59501	Adt59501	Plant pol
844	40	88.9	2464	4	ABB63174	Abb63174	Drosophil	917	39	86.7	131	9	ABM91613	Abm91613	M. xanthu
845	40	88.9	2566	8	ADS10740	Ads10740	Human the	918	39	86.7	135	5	ADK34119	Adk34119	Novel hum
846	40	88.9	3119	2	AAW36888	Aaw36888	Mouse Hun	919	39	86.7	136	4	AAO00750	Aao00750	Human pol
847	40	88.9	3129	2	AAW44743	Aaw44743	Mouse hun	920	39	86.7	139	3	AAV45011	Aav45011	Partial h
848	40	88.9	3139	2	AAV08898	Aav08898	Human Hun	921	39	86.7	140	3	AAV84588	Aav84588	A third p
849	40	88.9	3238	4	ABW71715	Abw71715	Drosophil	922	39	86.7	140	3	AAV68707	Aav68707	Amino aci
850	40	88.9	4019	4	AAE13839	Aae13839	Human lun	923	39	86.7	140	3	AAV68711	Aav68711	Amino aci
851	40	88.9	4019	7	ADP66733	Adp66733	Human lun	924	39	86.7	140	5	ABB82396	Abb82396	Human neu
852	40	88.9	4019	7	ADP66733	Adp66733	Human lun	924	39	86.7	140	5	ABB82396	Abb82396	Human neu
853	40	88.9	4025	5	ADP69736	Adp69736	Human pol	925	39	86.7	140	5	AAO22937	Aao22937	Human pre
854	40	88.9	4025	8	ADU18074	Adu18074	Human can	926	39	86.7	140	5	ABJ15110	Abj15110	Pre-pro-
855	40	88.9	19938	6	ABP76681	Abp76681	Human can	927	39	86.7	140	8	ADR16440	Adr16440	Neublasti
856	39	86.7	9	ADU03983	Adu03983	HTLV-I Ga	928	39	86.7	140	8	ADU46759	Adu46759	Human neu	
857	39	86.7	10	ADU03985	Adu03985	HTLV-I Ga	929	39	86.7	140	8	ADU46759	Adu46759	Human neu	
858	39	86.7	12	ADU03991	Adu03991	HTLV-I Ga	930	39	86.7	142	8	ADY72424	Ady72424	Plant ful	
859	39	86.7	13	ADU03987	Adu03987	HTLV-I Ga	931	39	86.7	147	8	ADY23420	Ady23420	Plant ful	
860	39	86.7	15	AAW38957	Aaw38957	Peptide r	932	39	86.7	150	5	ABP64467	Abp64467	Human ORF	
861	39	86.7	15	AAW38957	Aaw38957	Peptide r	933	39	86.7	153	7	ABM86340	Abm86340	Rice abio	
862	39	86.7	15	ADU03986	Adu03986	HTLV-I Ga	934	39	86.7	158	8	ADY23419	Ady23419	Plant ful	
863	39	86.7	15	ADT90816	Adt90816	PAGE2 prot	935	39	86.7	158	9	ADY35173	Ady35173	Signal pe	
864	39	86.7	16	AAW38947	Aaw38947	Peptide r	936	39	86.7	158	9	ABM92717	Abm92717	M. xanthu	
865	39	86.7	28	AAW39867	Aaw39867	E2 peptid	937	39	86.7	159	3	AAV44774	Aav44774	Protein-2	
866	39	86.7	34	ABO54390	Abos4390	Human gen	938	39	86.7	159	9	ADY35166	Ady35166	Immunoglo	
867	39	86.7	43	ADU04374	Adu04374	HTLV-I Ga	939	39	86.7	159	9	ADY35180	Ady35180	Signal pe	
868	39	86.7	50	AAV29804	Aav29804	Human GAB	940	39	86.7	161	3	AAG58353	Aag58353	Arabidops	
869	39	86.7	51	ADA48234	Ada48234	Rice prot	941	39	86.7	164	8	ABO60055	Abo60055	Human gen	
870	39	86.7	59	AAO23996	Aao23996	Human sec	942	39	86.7	165	2	AAV08459	Aav08459	Rat Ptx1	
871	39	86.7	61	AAO23996	Aao23996	Human sec	943	39	86.7	165	8	ADY79099	Ady79099	Plant ful	
872	39	86.7	62	ABO58404	Abos8404	Human gen	944	39	86.7	166	9	ADY35187	Ady35187	Signal pe	
873	39	86.7	67	AAW16772	Aaw16772	Peptide #	945	39	86.7	168	4	ABG13945	Abg13945	Novel hum	
874	39	86.7	67	ABB35758	Abb35758	Peptide #	946	39	86.7	171	3	AAV71561	Aav71561	Rat Nogo	
875	39	86.7	67	AAW29259	Aaw29259	Peptide #	947	39	86.7	178	4	AAW25955	Aaw25955	Human pro	
876	39	86.7	67	ABR30592	Abbr30592	Peptide #	948	39	86.7	178	7	ADC33057	Adc33057	Human nov	
877	39	86.7	67	ABR21182	Abbr21182	Protein #	949	39	86.7	178	8	ADM87723	Adm87723	Human EST	
878	39	86.7	67	AAW68948	Aaw68948	Human bon	950	39	86.7	179	9	ADY35160	Ady35160	Signal pe	
879	39	86.7	67	AAW56566	Aaw56566	Human bra	951	39	86.7	181	9	ADM45341	Adm45341	Human art	
880	39	86.7	67	ABG50611	Abg50611	Human liv	952	39	86.7	181	9	ADM45344	Adm45344	Human art	
881	39	86.7	67	AAW04489	Aaw04489	Peptide #	953	39	86.7	181	9	ADM45340	Adm45340	Human art	
882	39	86.7	67	ABG38532	Abg38532	Human pep	954	39	86.7	181	9	ADM45345	Adm45345	Human art	
883	39	86.7	68	AAV84596	Aav84596	Fragment	955	39	86.7	181	9	ADM45343	Adm45343	Human art	
884	39	86.7	68	AAV53874	Aav53874	Arabidops	956	39	86.7	181	9	ADM45342	Adm45342	Human art	
885	39	86.7	69	ADC32722	Adc32722	Human nov	957	39	86.7	185	8	ADQ16418	Adq16418	Nucleotid	
886	39	86.7	70	AAV10803	Aav10803	Arabidops	958	39	86.7	188	4	AAG82812	Aag82812	S. epider	
887	39	86.7	72	AAW14429	Aaw14429	HTLV-I Ga	959	39	86.7	188	7	ABM86285	Abm86285	Rice abio	
888	39	86.7	74	AAV53873	Aav53873	Arabidops	960	39	86.7	188	5	ABP38468	Abp38468	Staphyloc	
889	39	86.7	75	AAO05005	Aao05005	Human pol	961	39	86.7	198	8	ADU04629	Adu04629	Human vas	
890	39	86.7	76	AAV10802	Aav10802	Arabidops	962	39	86.7	202	7	AAO24273	Aao24273	Human vas	
891	39	86.7	77	ADV76639	Adv76639	Human TER	963	39	86.7	205	6	ABU20105	Abu20105	Protein e	
892	39	86.7	78	AAV53872	Aav53872	Arabidops	964	39	86.7	209	2	AAW26615	Aaw26615	Ferredoxi	
893	39	86.7	83	AAW95449	Aaw95449	Human rep	965	39	86.7	211	7	ADT88972	Adt88972	Tobacco M	
894	39	86.7	83	ABP09336	Abp09336	Human ORF	966	39	86.7	211	7	ADT88972	Adt88972	Tobacco M	
895	39	86.7	90	ADX78974	Adx78974	Plant ful	967	39	86.7	213	5	ABP90623	Abp90623	Chlamydia	
896	39	86.7	95	AAW58354	Aaw58354	Arabidops	968	39	86.7	213	8	ADW94802	Adw94802	Plant ful	
897	39	86.7	99	ADT57219	Adt57219	Plant pol	969	39	86.7	215	8	ADY11803	Ady11803	Plant ful	
898	39	86.7	104	AAW13657	Aaw13657	Peptide #	970	39	86.7	216	2	AAV35299	Aav35299	Chlamydia	
899	39	86.7	104	ABR32587	Abrr32587	Peptide #	971	39	86.7	220	3	AAV84583	Aav84583	Amino aci	
900	39	86.7	104	AAW26054	Aaw26054	Peptide #	972	39	86.7	220	3	AAV44776	Aav44776	Short spl	
901	39	86.7	104	ABW27438	Abw27438	Human pep	973	39	86.7	220	3	AAV44776	Aav44776	Short spl	

974 39 86.7 220 3 AAY68710 A human p
 975 39 86.7 220 4 AAB50978 Human PRO
 976 39 86.7 220 5 AAU86158 Human PRO
 977 39 86.7 220 5 AAB84975 Human PRO
 978 39 86.7 220 5 ABG30698 Human art
 979 39 86.7 220 5 ABB82388 Human neu
 980 39 86.7 220 5 ABB95581 Human ang
 981 39 86.7 220 5 AAO22940 Human foe
 982 39 86.7 220 6 ABUS6702 Lung canc
 983 39 86.7 220 6 ABUS6539 Lung canc
 984 39 86.7 220 6 ABUS6703 Lung canc
 985 39 86.7 220 6 ABUS6540 Lung canc
 986 39 86.7 220 6 ABU71444 Human neo
 987 39 86.7 220 7 ADD10607 Human sec
 988 39 86.7 220 7 ADD11567 Human sec
 989 39 86.7 220 7 ADJ37360 Human sec
 990 39 86.7 220 7 ADJ37343 Human tum
 991 39 86.7 220 7 ADN39086 Cancer/an
 992 39 86.7 220 7 ADN39084 Cancer/an
 993 39 86.7 220 8 ADK41568 Human sec
 994 39 86.7 220 8 ADG68267 Human PRO
 995 39 86.7 220 8 ADH43751 Human PRO
 996 39 86.7 220 8 ADK83096 Human PRO
 997 39 86.7 220 8 ADR16439 Human pre
 998 39 86.7 220 8 ADU46758 Human pre
 999 39 86.7 220 9 ADV35145 Human pre
 1000 39 86.7 220 9 ADZ80734 Amino aci

ALIGNMENTS

RESULT 1
 ID AAB17249 standard; peptide; 7 AA.
 AC AAB17249;
 XX
 DT 31-OCT-2000 (first entry)
 XX
 DE SH3 antagonist peptide sequence SEQ ID NO:305.
 XX
 KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
 KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
 KW immunosuppressive; EPO; TPO; CRLA4; mimetic; IL-1; TNF; antagonist; MMP;
 KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;
 KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
 KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
 KW thrombosis; pharmaceutical.
 XX
 OS Synthetic.
 XX
 PN WO200024782-A2.
 XX
 PD 04-MAY-2000.
 XX
 PF 25-OCT-1999; 99WO-US025044.
 XX
 PR 23-OCT-1998; 98US-0105371P.
 PR 22-OCT-1999; 99US-00428082.
 XX
 PA (AMGE-) AMGEN INC.
 XX
 PI Feige U, Liu C, Cheetham J, Boone TC;
 XX
 DR WPI; 2000-350702/30.
 XX
 XX Novel composition of matter comprising an Fc domain and pharmacologically
 PT active peptides, useful for treating cancer and autoimmune diseases.
 XX
 PS Claim 39; Page 302; 608pp; English.
 XX
 XX The present invention describes composition of matter (I) comprising an

CC domain, pharmacologically active peptides, and linkers. Where (I) is:
 (X1)a-P1-(X2)b, where: P1 = an Fc domain; X1 and X2 = are each
 independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
 (L2)d-P2-(L3)e-P3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
 P3, and P4 = are each independently sequences of pharmacologically active
 peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,
 c, d, e, and f = are each independently 0 or 1, provided that at least 1
 of a and b is 1. The composition can have cytostatic, antiasthmatic,
 thrombolytic and immunosuppressive activities. DNAs, vectors and host
 cells from the present invention can be used for producing pharmaceutical
 compositions. The compositions are useful for treating cancer, asthma,
 thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
 a Fab domain) can provide a longer half-life or incorporate functions
 such as Fc receptor binding, protein A binding, complement fixation, and
 possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to
 AAB18003 represent nucleotide and amino acid sequences used in the
 exemplification of the present invention

Query Match 100.0%; Score 45; DB 3; Length 7;
 Best Local Similarity 100.0%; Pred. No. 2e+06;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 RPQPPPP 7
 DB 1 RPQPPPP 7

RESULT 2
 ABB73242
 ID ABB73242 standard; peptide; 7 AA.
 AC ABB73242;
 XX
 DT 05-APR-2002 (first entry)
 XX
 DE Src homology3 (SH3) antagonist peptide SEQ ID NO:305.
 XX
 KW Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
 KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
 KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;
 KW TPO mimetic peptide; EPO mimetic peptide; BMP; VEGF antagonist;
 KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
 KW cycostatic; antirheumatic; antiarthritis; antidiabetic; ophthalmological;
 KW anianaemic; anorectic; antiinfertility; haemostatic; dermatological;
 KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
 KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
 KW sleep disorder; neurological degenerative disease; anaemia;
 KW thrombocytopenia; metastatic tumour; systemic lupus erythematosus;
 KW Fanconi's syndrome.
 XX
 OS Homo sapiens.
 OS Synthetic.
 XX
 PN WO200183525-A2.
 XX
 PD 08-NOV-2001.
 XX
 PF 02-MAY-2001; 2001WO-US014310.
 XX
 PR 03-MAY-2000; 2000US-00563286.
 XX
 PA (AMGE-) AMGEN INC.
 XX
 PI Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;
 XX
 DR WPI; 2002-130313/17.
 XX
 XX Novel vehicle-peptide molecule or its multimers useful for treating
 PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
 PT diabetic retinopathy, obesity, sleep disorders and infertility.
 XX

PS Claim 39; Page 55; 176pp; English.

XX The present invention describes a vehicle-peptide molecule (I) or its
 CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,
 CC cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,
 CC antianemic, anorectic, antifertility, haemostatic, dermatological and
 CC neuroprotective activities. (I) can be used as a therapeutic or
 CC prophylactic agent as well as for screening purposes. (I) is useful for
 CC diagnosing diseases characterised by dysfunction of their associated
 CC protein of interest, for identifying normal or abnormal proteins of
 CC interest, as a part of diagnostic kit to detect the presence of their
 CC proteins of interest in a biological sample. Additionally, (I) is useful
 CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,
 CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders, EPO-
 CC infertility, and neurological degenerative diseases. (I), comprising EPO-
 CC mimetic compounds are useful for treating disorders characterised by low
 CC red blood cell levels such as anaemia. The TPO-mimetic comprising
 CC compounds are useful for treating conditions that involve an existing
 CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet
 CC deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic
 CC tumour which result in thrombocytopaenia, systemic lupus erythematosus,
 CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777
 CC represent amino acid and nucleic acid sequences used in the
 CC exemplification of the present invention

XX
 CC Sequence 7 AA;

Query Match 100.0%; Score 45; DB 5; Length 7;
 Best Local Similarity 100.0%; Pred. No. 2e+06;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
 Db ||||| 7
 1 RPQPPPP 7

RESULT 3
 ADJ73396
 ID ADJ73396 standard; peptide; 7 AA.
 XX
 AC ADJ73396;
 XX
 DT 06-MAY-2004 (first entry)
 DE SH3 antagonist peptide sequence SeqID 851.
 XX
 KW mimetic; CDR mimetibody; gene therapy; transgenic; immune;
 KW cardiovascular; infectious; malignant; neurologic disease; anaemia;
 KW immunomodulator; cardiant; antimicrobial; cytostatic; neuroprotective;
 KW SH3.
 XX
 OS Synthetic.
 XX
 PN WO2003084477-A2.
 XX
 PD 16-OCT-2003.
 XX
 PF 24-MAR-2003; 2003WO-US009139.
 XX
 PR 29-MAR-2002; 2002US-036879P.
 XX
 PA (CENZ) CENTOCOR INC.
 XX
 PI Heavner GA, Knight DM, Scallion BJ, Ghayeb J;
 XX
 DR WPI; 2003-804237/75.
 XX
 PT New CDR mimetibody comprising a portion of a heavy or light chain
 PT variable region comprising human framework or ligand binding region,
 PT useful for preparing a composition for treating e.g., immune,
 PT cardiovascular or neurologic disease.
 XX
 PS Disclosure; SEQ ID NO 851; 97pp; English.

XX This invention relates to novel mammalian CDR mimetibodies, specific
 CC portions or variants thereof. Specifically, it refers to an antibody
 CC fragment where a protein has been inserted into, or replaces a portion
 CC of, one or more CDR regions, such that each CDR mimetibody comprises at
 CC least one portion of a heavy chain or light chain variable region, which
 CC itself comprises at least one human framework region and at least one
 CC ligand binding region (LBR). The present invention describes human
 CC mimetibodies, including modified immunoglobulins and cleavage products
 CC that can be useful in gene therapy and the generation of transgenic
 CC plants and animals. Furthermore, the CDR mimetibody is useful for
 CC preparing compositions for modulating, treating or reducing the symptoms
 CC of immune, cardiovascular, infectious, malignant and/ or neurologic
 CC diseases, as well as anaemia. Accordingly, they exhibit immunomodulator,
 CC cardiant, antimicrobial, cytostatic and neuroprotective activities. This
 CC peptide sequence is an SH3 antagonist peptide sequence used to make a
 CC mimetibody of the invention.

XX
 CC Sequence 7 AA;

Query Match 100.0%; Score 45; DB 7; Length 7;
 Best Local Similarity 100.0%; Pred. No. 2e+06;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
 Db ||||| 7
 1 RPQPPPP 7

RESULT 4
 ADJ53030
 ID ADJ53030 standard; peptide; 7 AA.
 XX
 AC ADJ53030;
 XX
 DT 06-MAY-2004 (first entry)
 DE CH1 deleted mimetibody-related peptide SeqID851.
 XX
 KW CH1 deleted mimetibody; immunosuppressive; cardiovascular; cardiant;
 KW hypotensive; neuroprotective; nootropic; antibacterial; virucide;
 KW fungicide; gene therapy; immune disorder; cardiovascular disease;
 KW arrhythmia; hypertension; heart failure; neurodegenerative;
 KW multiple sclerosis; dementia; Alzheimer's disease; anaemia;
 KW cancerous condition; infectious disease; bacterial infection;
 KW viral infection; fungal infection.
 XX
 OS Unidentified.
 OS Synthetic.
 XX
 PN WO2004002417-A2.
 XX
 PD 08-JAN-2004.
 XX
 PF 27-JUN-2003; 2003WO-US020347.
 XX
 PR 28-JUN-2002; 2002US-0392431P.
 XX
 PA (CENZ) CENTOCOR INC.
 XX
 PI Heavner GA, Knight DM, Ghayeb J, Scallion BJ, Nesspor TC;
 PI Kutoloski KA;
 XX
 DR WPI; 2004-082870/08.
 XX
 PT New CH1-deleted mimetibody polypeptides and nucleic acids, useful for
 PT modulating, treating, alleviating, preventing an immune, cardiovascular,
 PT or neurodegenerative disease or disorder, anemia, cancer, or infectious
 PT diseases.
 XX
 PS Claim 3; SEQ ID NO 851; 129pp; English.
 XX
 CC This invention relates to CH1 deleted mimetibodies (and the DNA sequences

CC which encode them), compositions, methods and uses. The invention may be
 CC useful for the development of compounds with an immunosuppressive,
 CC cardiovascular, cardiac, hypotensive, neuroprotective, nootropic,
 CC antibacterial, virucide or fungicide activity. In addition, the disclosed
 CC sequences may prove useful for gene therapy. The CHI-deleted mimetibody
 CC is useful for diagnosing or treating a disease condition in a cell,
 CC tissue, organ or animal, specifically for modulating, treating,
 CC alleviating, preventing the incidence or reducing the symptoms of an
 CC immune, cardiovascular (for example arrhythmia, hypertension or heart
 CC failure), or neurodegenerative (for example multiple sclerosis, dementia
 CC or Alzheimer's disease) diseases or disorders, anaemia, cancerous
 CC conditions, or infectious diseases (for example bacterial, viral or
 CC fungal infection). The present sequence is that of a peptide which may be
 CC used during the creation of a mimetibody of the invention.

XX Sequence 7 AA;

Query Match 100.0%; Score 45; DB 8; Length 7;
 Best Local Similarity 100.0%; Pred. No. 2e+06;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 RPQPPPP 7
 Db 1 RPQPPPP 7

RESULT 5
 ADJ51991
 ID ADJ51991 standard; peptide; 7 AA.

AC ADJ51991;

DT 06-MAY-2004 (first entry)

DE CHI deleted mimetibody-related peptide SeqID851.

KW CHI deleted mimetibody; osteopathic; cardiovascular-Gen;
 KW dermatological-Gen; auditory; endocrine-Gen; gastrointestinal-Gen;
 KW gynaecological-Gen; hepatotropic; haemostatic; immunomodulator;
 KW anti-allergic; muscular-Gen; cytostatic; anti-inflammatory; neuroleptic;
 KW ophthalmological; nephrotropic; respiratory-Gen; tumour necrosis factor;
 KW TNF; cytokine; bone disorder; joint disorder; cardiovascular disorder;
 KW dental disorder; oral disorder; dermatological disorder; ear disorder;
 KW nose disorder; throat disorder; endocrine disorder; metabolic disorder;
 KW gastrointestinal disorder; gynaecological disorder; hepatic disorder;
 KW obstetric disorder; haematologic disorder; immunological disorder;
 KW allergic disorder; infectious disorder; musculoskeletal disorder;
 KW oncological disorder; neurological disorder; nutritional disorder;
 KW ophthalmologic disorder; pediatric disorder; psychiatric disorder;
 KW renal disorder; pulmonary disorder.

XX Unidentified.

OS Synthetic.

PN WO2004002424-A2.

PD 08-JAN-2004.

PF 30-JUN-2003; 2003WO-US020495.

PR 28-JUN-2002; 2002US-0392431P.

PR 19-SEP-2002; 2002US-0412144P.

XX (CENZ) CENTOCOR INC.

XX Heavner GA, Knight DM, Ghayeb J, Scallon BJ, Nessapor TC;

PI Kutloski KA;

XX WPI; 2004-082872/08.

XX New CHI deleted mimetibody polypeptide and nucleic acid, useful for
 PT diagnosing, preventing or treating cardiovascular, dermatologic,
 PT endocrine, gastrointestinal, gynecologic, infectious, neurologic and

PT nutritional disorders.

XX Claim 15; SEQ ID NO 851; 123pp; English.

XX This invention relates to CHI deleted mimetibodies (and the DNA sequences
 CC which encode them), compositions, methods and uses. The invention may be
 CC useful for the development of compounds with an osteopathic,
 CC cardiovascular-Gen, dermatological-Gen, auditory, endocrine-Gen,
 CC gastrointestinal-Gen, gynaecological-Gen, hepatotropic, haemostatic,
 CC immunomodulator, anti-allergic, ophthalmological, nephrotropic or
 CC anti-inflammatory, neuroleptic, musculo-Gen, cytostatic,
 CC respiratory-Gen activity acting as a tumour necrosis factor (TNF) -
 CC modulator or cytokine-agonist. The methods and compositions of the
 CC present invention are useful for the diagnosis, prevention and/or
 CC treatment of diseases or conditions associated with aberrant expression
 CC or activity of the CHI deleted mimetibody, such as a bone or joint,
 CC cardiovascular, dental or oral, dermatological, ear, nose or throat,
 CC endocrine, metabolic, gastrointestinal, gynaecological, hepatic,
 CC obstetric, haematologic, immunological, allergic, infectious,
 CC musculoskeletal, oncological, neurological, nutritional, ophthalmologic,
 CC pediatric, psychiatric, renal or pulmonary disorders. The present
 CC sequence is that of a peptide which may be used during the creation of a
 CC mimetibody of the invention.

XX Sequence 7 AA;

Query Match 100.0%; Score 45; DB 8; Length 7;
 Best Local Similarity 100.0%; Pred. No. 2e+06;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 RPQPPPP 7
 Db 1 RPQPPPP 7

RESULT 6

AAW11113

ID AAW11113 standard; peptide; 13 AA.

AC AAW11113;

DT 25-JUN-1997 (first entry)

DE Src SH3 domain-binding peptide used in signal transduction modulation.

XX Src; SH3; Src homology region 3; binding affinity; oncogenic protein;
 KW protein tyrosine kinase; signal transduction; RNA processing;
 KW trafficking; translation.

XX Synthetic.

XX WO9603649-A1.

PD 08-FEB-1996.

PF 24-JUL-1995; 95WO-US009382.

PR 22-JUL-1994; 94US-00278865.

PR 07-JUN-1995; 95US-00483555.

XX (UYNC-) UNIV NORTH CAROLINA.

XX Sparks AB, Kay BK, Thorn JM, Quilliam LA, Der CJ;

XX WPI; 1996-117151/12.

XX Peptide with binding affinity for Src homology region 3 (SH3) domains of
 PT proteins - useful for e.g. modulating signal transduction pathways at the
 PT cellular level, esp. protein tyrosine kinase-mediated.

XX Claim 38; Page 87; 116pp; English.

XX AAW11098-W11124 are peptides that bind to the Src SH3 domain. The SH3

CC binding peptides are useful in modulating signal transduction pathways at
 CC the cellular level (especially protein tyrosine kinase-mediated),
 CC modulating oncogenic protein activity, or providing compounds for the
 CC development of drugs with the ability to modulate broad classes, as well
 CC as specific classes, of proteins involved in signal transduction and also
 CC for regulating the processing, trafficking or translation of RNA.
 CC Conjugates of the peptides with detectable labels or imaging agents are
 CC useful for imaging cells, tissues and organs in which Src or Src-related
 CC proteins are expressed
 XX
 SQ Sequence 13 AA;

Query Match 100.0%; Score 45; DB 2; Length 13;
 Best Local Similarity 100.0%; Pred. No. 16;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
 Db 4 RPQPPPP 10
 |||||

RESULT 7

AAW2511
 ID AAW2511 standard; peptide; 31 AA.

XX AC AAW2511;

XX DT 27-MAR-1998 (first entry)

XX DE Random peptide recombinant clone R8C.YES3.9.

XX KW Cortactin; SH3 domain; binding peptide; Src homology region 3;
 KW tyrosine kinase; immune response; lymphokine; interleukin 1; Nck; Abl;
 KW PLCgamma; p53bp2; Crk; Yes; Grb2.

XX OS Synthetic.
 OS Unidentified.

XX PN WO9730074-A1.

XX PD 21-AUG-1997.

XX PF 14-FEB-1997; 97WO-US002298.

XX PR 16-FEB-1996; 96US-00602999.

XX PA (CYTO-) CYTOGEN CORP.

XX PA (UINC-) UNIV NORTH CAROLINA.

XX PI Sparks AB, Kay BK, Thorn JM, Quilliam LA, Der CJ, Fowlkes DM;
 PI Rider JE;

XX PP WPI; 1997-424972/39.

XX SR Src homology region 3 binding peptide - used to activate Src tyrosine
 PT kinase(s) and to stimulate immune response by increasing production of
 PT certain lymphokine(s), e.g. interleukin-1.

XX PS Disclosure; Fig 5; 131pp; English.

XX CC The present sequence represents a random peptide recombinant isolated by
 CC the method of the present invention. SH3 (Src homology region 3) binding
 CC peptides are selected from: (a) peptides which bind the SH3 domain of
 CC Cortactin; (b) peptides which bind the middle SH3 domain of Nck; (c)
 CC peptides which bind the SH3 domain of Abl; (d) peptides which bind the
 CC SH3 domain of Src; (e) peptides which bind the SH3 domain of PLC gamma;
 CC (f) peptides which bind the SH3 domain of p53bp2; (g) peptides which bind
 CC the amino-terminal SH3 domain of Crk; (h) peptides which bind the SH3
 CC domain of Yes; and (i) peptides which bind the amino-terminal SH3 domain
 CC of Grb2. The purified binding peptides can be used in the method to
 CC identify inhibitors of their binding to their respective SH3 domains,
 CC which could be used to modulate the pharmacological activity of proteins
 CC or polypeptide containing the SH3 domain. The peptides can also be used

CC to activate Src or Src-related protein tyrosine kinases, to stimulate the
 CC immune response by increasing the production of certain lymphokines, e.g.
 CC tumour necrosis factor-alpha and interleukin-1, or to deliver a
 CC conjugated molecule to certain cellular compartments containing Src or
 CC Src related proteins
 XX
 SQ Sequence 31 AA;

Query Match 100.0%; Score 45; DB 2; Length 31;
 Best Local Similarity 100.0%; Pred. No. 33;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
 Db 18 RPQPPPP 24
 |||||

RESULT 8

ADI45407

ID ADI45407 standard; protein; 122 AA.

XX AC ADI45407;

XX DT 22-APR-2004 (first entry)

XX DE Rice isoprenoid biosynthesis-associated protein #169.

XX KW Rice; isoprenoid biosynthesis; plant; isopentenyl diphosphate; IPP;
 KW dimethylallyl alcohol; DMAPP; short-chain plastid prenyltransferase;
 KW gibberellin; carotenoid; abscisic acid; tocopherol; plastoquinone;
 KW phyloquinone; mevalonate pathway; phytosterol; brassinosteroid;
 KW ubiquinone; monoterpene; sesquiterpene; protein prenylation; chlorophyll;
 KW haeme; yield.

XX OS Oryza sativa.

XX PN US2004010815-A1.

XX PD 15-JAN-2004.

XX PF 26-SEP-2002; 2002US-00259194.

XX PR 26-SEP-2001; 2001US-0325277P.

XX PR 04-APR-2002; 2002US-0370620P.

XX PR 04-APR-2002; 2002US-0370743P.

XX PA (LANG/) LANGE B M.

XX PA (GHAS/) GHASSEMIAN M.

XX PA (BRIG/) BRIGGS S P.

XX PA (COOP/) COOPER B.

XX PA (GLAZ/) GLAZEBROOK J.

XX PA (GOFF/) GOFF S A.

XX PA (KATA/) KATAGIRI F.

XX PA (KREP/) KREPS J.

XX PA (MOUG/) MOUGHAMER T.

XX PA (PROV/) PROVART N.

XX PA (RICK/) RICHE D.

XX PA (ZHUT/) ZHU T.

XX PI Lange BM, Ghassenian M, Briggs SP, Cooper B, Glazebrook J;

PI Goff SA, Katagiri F, Kreps J, Moughamer T, Provart N, Ricke D;
 PI Zhu T;

XX DR WPI; 2004-090562/09.

XX DR N-PSDB; ADI45406.

XX CC New isolated polynucleotides and polypeptides associated with isoprenoid
 PT synthesis in plants, useful for producing transgenic plants, for targeted
 PT gene disruption, as well as markers or probes.

XX PS Claim 4; SEQ ID NO 338; 117pp; English.

XX CC The invention relates to a polynucleotide (or its complement, protein

CC encoding fragment or reverse complement), comprising a nucleotide
 CC sequence encoding a polypeptide comprising an amino acid sequence
 CC involved in or associated with the biosynthesis of isoprenoids in a rice
 CC plant. Also included are an isolated polypeptide involved in or
 CC associated with the biosynthesis of isoprenoids in a plant, an expression
 CC cassette comprising the polynucleotide, a host cell comprising the
 CC expression cassette, and a transgenic plant comprising the expression
 CC cassette. The polypeptides and polynucleotides include those associated
 CC with the biosynthesis of isopentenyl diphosphate (IPP) and dimethylallyl
 CC alcohol (DMAPP), the biosynthesis of short-chain plastid
 CC prenyltransferases, the biosynthesis of gibberellins, the biosynthesis of
 CC carotenoids and/or abscisic acids, the biosynthesis of tocopherols,
 CC plastoquinone and/or phyloquinone biosynthesis, the mevalonate pathway,
 CC phytosterol and brassinosteroid metabolism, biosynthesis of ubiquinone,
 CC biosynthesis of monoterpenes and sesquiterpenes, protein prenylation, and
 CC biosynthesis of chlorophyll or haeme. Also disclosed are banana, wheat
 CC and corn homologues of some of the rice polynucleotides. The
 CC polynucleotides are useful for producing transgenic plants, where the
 CC genome is augmented by a nucleic acid molecule of the invention, or in
 CC which the corresponding gene has been disrupted, e.g. to result in a
 CC loss, a decrease or an alteration in the function of the product encoded
 CC by the gene. The plants may also have increased yields and/or produce a
 CC better quality product than the corresponding wild-type plant. The
 CC nucleic acid molecules are useful for targeted gene disruption, as well
 CC as markers and probes. Note: The sequence data for this patent did not
 CC form part of the printed specification, but was obtained in electronic
 CC format directly from USPTO at
 CC seqdata.uspto.gov/sequence.html?DocID=20040010815. The present sequence
 CC represents a Rice isoprenoid biosynthesis-associated protein of the
 CC invention.

XX SQ Sequence 122 AA;

Query Match 100.0%; Score 45; DB 8; Length 122;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
 DB 113 RPQPPPP 119

RESULT 9
 ABB97921
 ID ABB97921 standard; protein; 173 AA.

XX AC ABB97921;

XX DT 03-OCT-2002 (first entry)

XX DE Human secretory polypeptide (SPTM) 173.

XX KW Human; secretory protein; secretory polynucleotides; SPTM;
 KW SPTM-related disease; somatic gene therapy; germline gene therapy;
 KW severe combined immunodeficiency; intracellular parasite protection;
 KW fungal parasite; protozoan parasite; cell proliferative disorder; cancer;
 KW immune disorder; AIDS; neurological disorder; Parkinson's disease;
 KW motor neuron disorder; demyelinating disease; multiple sclerosis;
 KW meningitis; abscess; prion diseases; cerebral palsy;
 KW neuroskeletal disorder; peripheral nervous system disorder;
 KW dermatomyositis; polymyositis; myopathy; myasthenia gravis;
 KW mental disorder; Tourette's syndrome.

XX OS Homo sapiens.

XX PN WO200220756-A2.

XX PD 14-MAR-2002.

XX PF 30-AUG-2001; 2001WO-US027297.

XX PR 05-SEP-2000; 2000US-0229747P.

PR 05-SEP-2000; 2000US-0229748P.

PR 05-SEP-2000; 2000US-0229749P.
 PR 05-SEP-2000; 2000US-0229750P.
 PR 05-SEP-2000; 2000US-0229751P.
 PR 05-SEP-2000; 2000US-0230016P.
 PR 05-SEP-2000; 2000US-0230583P.
 PR 06-SEP-2000; 2000US-0230505P.
 PR 06-SEP-2000; 2000US-0230514P.
 PR 06-SEP-2000; 2000US-0230515P.
 PR 06-SEP-2000; 2000US-0230517P.
 PR 06-SEP-2000; 2000US-0230518P.
 PR 06-SEP-2000; 2000US-0230519P.
 PR 06-SEP-2000; 2000US-0230595P.
 PR 06-SEP-2000; 2000US-0230596P.
 PR 06-SEP-2000; 2000US-0230597P.
 PR 06-SEP-2000; 2000US-0230599P.
 PR 06-SEP-2000; 2000US-0230610P.
 PR 06-SEP-2000; 2000US-0230864P.
 PR 06-SEP-2000; 2000US-0230865P.
 PR 06-SEP-2000; 2000US-0230988P.
 PR 06-SEP-2000; 2000US-0230989P.
 PR 06-SEP-2000; 2000US-0230990P.
 PR 07-SEP-2000; 2000US-0230896P.
 PR 07-SEP-2000; 2000US-0230897P.
 PR 07-SEP-2000; 2000US-0230951P.
 PR 07-SEP-2000; 2000US-0231163P.
 PR 07-SEP-2000; 2000US-0231832P.

(INCY-) INCYTE GENOMICS INC.

XX PA Stuart J, Lincoln SE, Altus CM, Dufour GE, Chalup MS, Hillman JL;
 XX PI Jones AL, Yu JY, Wright RJ, Gietzen D, Liu TF, Yap PE, Dahl CR;
 XX PI Moniyama MG, Bradley DL, Rohatgi SD, Harris B, Roseberry AM;
 XX PI Gerstlin EH, Peralta CH, David MH, Panzer SR, Flores V, Daifo A;
 XX PI Marwaha R, Chen AJ, Chang SC, Au AP, Inman RR;
 XX DR WPI; 2002-315658/35.
 XX N-PSDB; ABL99918.

PT Polynucleotide sequences encoding human secretory proteins useful for
 PT gene therapy of e.g. genetic deficiency disorders, cancers, and diseases
 PT caused by intracellular parasites.

XX Claim 29; Page 479; 585pp; English.

XX The invention comprises the amino acid and coding sequences of human
 CC secretory (SPTM) proteins. The SPTM DNA and amino acid sequences are
 CC useful for treating a disease or condition associated with the expression
 CC of functional SPTM. The SPTM DNA sequences are useful for somatic or
 CC germline gene therapy to correct a genetic deficiency (e.g. severe
 CC combined immunodeficiency). The SPTM DNA sequences are also useful in
 CC providing protection against intracellular parasites (e.g. fungal
 CC parasites and protozoan parasites). The SPTM DNA and protein sequences
 CC are also useful for diagnosing cell proliferative disorders, cancer,
 CC immune disorders (e.g. AIDS), neurological disorders (e.g. Parkinson's
 CC disease), motor neuron disorders, demyelinating diseases (e.g. multiple
 CC sclerosis), meningitis, abscesses, prion diseases, cerebral palsy,
 CC neuroskeletal disorders, peripheral nervous system disorders,
 CC dermatomyositis and polymyositis, myopathy, myasthenia gravis, and mental
 CC disorders (e.g. Tourette's syndrome). Amino acid sequences ABB97749 -
 CC ABB97933 represent human secretory proteins of the invention

XX SQ Sequence 173 AA;

Query Match 100.0%; Score 45; DB 5; Length 173;

Best Local Similarity 100.0%; Pred. No. 1.4e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7

DB 106 RPQPPPP 112

RESULT 10

AB072442
ID AB072442 standard; protein; 249 AA.
XX
AC AB072442;
XX
DT 29-JUL-2004 (first entry)
XX
DE Pseudomonas aeruginosa polypeptide #4617.
XX
KW Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
XX
OS Pseudomonas aeruginosa.
XX
PN US6551795-B1.
XX
PD 22-APR-2003.
XX
PF 18-FEB-1999; 99US-00252991.
XX
PR 18-FEB-1998; 98US-0074788P.
PR 27-JUL-1998; 98US-0094190P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;
XX
DR WPI; 2003-615309/58.
DR N-PSDB; ABD06013.
XX
PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
PT pathological conditions resulting from bacterial infection.
XX
PS Disclosure; SEQ ID NO 21188; 455pp; English.
XX
CC The invention relates to Pseudomonas aeruginosa polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
CC for the ability to bind a P. aeruginosa nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-P. aeruginosa drugs, as templates for recombinant
CC production of P. aeruginosa-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of P. aeruginosa-caused
CC infection, and in detection of P. aeruginosa sequences or other sequences
CC of Pseudomonas species using biochip technology. Sequences ABO67826-
CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html
SQ Sequence 249 AA;
XX
Query Match 100.0%; Score 45; DB 7; Length 249;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RQPPPPP 7
Db 43 RQPPPPP 49
RESULT 11
ADP29798
ID ADP29798 standard; protein; 276 AA.
XX
AC ADP29798;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human secreted protein SEQ ID #565.
XX

KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
XX cancer; inflammatory; immune; human secreted protein.
OS Homo sapiens.
XX
PN WO2004035732-A2.
XX
PD 29-APR-2004.
XX
PF 28-AUG-2003; 2003WO-US026780.
XX
PR 29-AUG-2002; 2002US-0406576P.
PR 29-AUG-2002; 2002US-0406579P.
PR 29-AUG-2002; 2002US-0406585P.
PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 22-MAY-2003; 2003US-0472402P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.

```
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
XX Williams JT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Huang MM, Kothakota S, Haishan L, Linnemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX
XX MPI; 2004-348438/32.
XX
XX New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
XX Claim 1; SEQ ID NO 1796; 428pp; English.
XX
XX The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antitumour, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPWEB and is not in the specification.
XX
XX Sequence 276 AA;
SQ
Query Match 100.0%; Score 45; DB 8; Length 276;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RPQPPPP 7
DB 188 RPQPPPP 194
RESULT 12
AAG13899
ID AAG13899 standard; protein; 288 AA.
XX
XX AAG13899;
AC
XX
XX 17-OCT-2000 (first entry)
DT
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 13560.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
XX Arabidopsis thaliana.
OS
XX
XX EP1033405-A2.
PN
XX
XX 06-SEP-2000.
PD
XX
XX 25-FEB-2000; 2000EP-00301439.
PF
XX
XX 25-FEB-1999; 99US-0121825P.
PR
XX 05-MAR-1999; 99US-0123180P.
PR
XX 09-MAR-1999; 99US-0123548P.
PR
XX 23-MAR-1999; 99US-0125788P.
PR
XX 29-MAR-1999; 99US-0126264P.
PR
XX 29-MAR-1999; 99US-0126785P.
PR
XX 01-APR-1999; 99US-0127462P.
PR
XX 06-APR-1999; 99US-0128234P.
PR
XX 08-APR-1999; 99US-0128714P.
PR
XX 16-APR-1999; 99US-0129845P.
PR
19-APR-1999; 99US-0130077P.
PR
21-APR-1999; 99US-0130449P.
PR
23-APR-1999; 99US-0130510P.
PR
23-APR-1999; 99US-0130891P.
PR
30-APR-1999; 99US-0131449P.
PR
30-APR-1999; 99US-0132048P.
PR
30-APR-1999; 99US-0132407P.
PR
04-MAY-1999; 99US-0132484P.
PR
05-MAY-1999; 99US-0132485P.
PR
06-MAY-1999; 99US-0132486P.
PR
06-MAY-1999; 99US-0132487P.
PR
07-MAY-1999; 99US-0132863P.
PR
11-MAY-1999; 99US-0134256P.
PR
14-MAY-1999; 99US-0134218P.
PR
14-MAY-1999; 99US-0134219P.
PR
14-MAY-1999; 99US-0134221P.
PR
14-MAY-1999; 99US-0134370P.
PR
18-MAY-1999; 99US-0134768P.
PR
19-MAY-1999; 99US-0134941P.
PR
20-MAY-1999; 99US-0135124P.
PR
21-MAY-1999; 99US-0135353P.
PR
24-MAY-1999; 99US-0135629P.
PR
25-MAY-1999; 99US-0136021P.
PR
27-MAY-1999; 99US-0136392P.
PR
28-MAY-1999; 99US-0136782P.
PR
01-JUN-1999; 99US-0137222P.
PR
03-JUN-1999; 99US-0137528P.
PR
04-JUN-1999; 99US-0137502P.
PR
07-JUN-1999; 99US-0137724P.
PR
08-JUN-1999; 99US-0138094P.
PR
10-JUN-1999; 99US-0138540P.
PR
10-JUN-1999; 99US-0138847P.
PR
14-JUN-1999; 99US-0139119P.
PR
16-JUN-1999; 99US-0139452P.
PR
16-JUN-1999; 99US-0139453P.
PR
17-JUN-1999; 99US-0139492P.
PR
18-JUN-1999; 99US-0139454P.
PR
18-JUN-1999; 99US-0139455P.
PR
18-JUN-1999; 99US-0139456P.
PR
18-JUN-1999; 99US-0139457P.
PR
18-JUN-1999; 99US-0139763P.
PR
18-JUN-1999; 99US-0139763P.
PR
21-JUN-1999; 99US-0139817P.
PR
22-JUN-1999; 99US-0139899P.
PR
23-JUN-1999; 99US-0140353P.
PR
23-JUN-1999; 99US-0140354P.
PR
24-JUN-1999; 99US-0140695P.
PR
28-JUN-1999; 99US-0140823P.
PR
29-JUN-1999; 99US-0140991P.
PR
30-JUN-1999; 99US-0141287P.
PR
01-JUL-1999; 99US-0141842P.
PR
01-JUL-1999; 99US-0142154P.
PR
02-JUL-1999; 99US-0142055P.
PR
06-JUL-1999; 99US-0142390P.
PR
08-JUL-1999; 99US-0142803P.
PR
09-JUL-1999; 99US-0142920P.
PR
12-JUL-1999; 99US-0142977P.
PR
13-JUL-1999; 99US-0143542P.
PR
14-JUL-1999; 99US-0143624P.
PR
15-JUL-1999; 99US-0144005P.
PR
16-JUL-1999; 99US-0144085P.
PR
16-JUL-1999; 99US-0144086P.
PR
19-JUL-1999; 99US-0144325P.
PR
19-JUL-1999; 99US-0144331P.
PR
19-JUL-1999; 99US-0144332P.
PR
19-JUL-1999; 99US-0144333P.
PR
19-JUL-1999; 99US-0144334P.
```

PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 21-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145911P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149317P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151338P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.

PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 100.0%; Score 45; DB 3; Length 288;

Best Local Similarity 100.0%; Pred. No. 2.2e+02; Mismatches 0; Indels 0; Gaps 0;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7

Db 53 RPQPPPP 59

RESULT 13

ABM97421

ID ABM97421 standard; protein; 314 AA.

AC ABM97421;

XX

DT 02-JUN-2005 (first entry)

DE M. xanthus protein sequence, seq id 16620.

XX Transgenic plant; DNA replication; gene regulation; gene expression.

OS Myxococcus xanthus.

PN US6833447-B1.

PD 21-DEC-2004.

PP 10-JUL-2001; 2001US-00902540.

PR 10-JUL-2000; 2000US-0217883P.

XX (MONS) MONSANTO TECHNOLOGY LLC.

PI Goldman BS, Hinkle GJ, Slater SC, Wiegand RC;

DR WPI; 2005-028716/03.

XX New substantially purified Myxococcus xanthus nucleic acid molecule

PT encoding a nitrite reductase, useful for determining gene expression,

PT identifying mutations in a gene of interest, and for constructing

XX mutations in a gene of interest.

PS Example 2; SEQ ID NO 16620; 25pp; English.

XX The invention relates to a substantially purified nucleic acid molecule

CC encoding a nitrite reductase of SEQ ID NO. 11926. Further disclosed is a

CC recombinant DNA construct for expression of a nitrite reductase gene in a

CC plant cell, and a plant cell comprising the recombinant DNA construct.

CC The nucleic acid is useful for determining gene expression, identifying
 CC mutations in a gene of interest, and for constructing mutations in a gene
 CC of interest. Sequences given in records for SEQ IDs 9692-16825 represent
 CC a group of 7134 *Mycococcus xanthus* proteins and peptides. Note: The
 CC sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from USPTO

XX Sequence 314 AA;

Query Match 100.0%; Score 45; DB 9; Length 314;
 Best Local Similarity 100.0%; Pred. No. 2.4e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQPPPP 7
 DB 8 RQPPPP 14

RESULT 14

AGL13898 ID AAG13898 standard; protein; 322 AA.

AC AAG13898;

DT 17-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 13559.

DE Arabidopsis thaliana protein fragment SEQ ID NO: 13559.
 XX Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridization assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.

XX Arabidopsis thaliana.

XX EP1033405-A2.

PD 06-SEP-2000.

XX 25-FEB-2000; 2000EP-00301439.

PR 25-FEB-1999; 99US-0121825P.

PR 05-MAR-1999; 99US-012180P.

PR 09-MAR-1999; 99US-0121548P.

PR 23-MAR-1999; 99US-0125788P.

PR 25-MAR-1999; 99US-0126264P.

PR 29-MAR-1999; 99US-0126785P.

PR 01-APR-1999; 99US-0127462P.

PR 06-APR-1999; 99US-0128234P.

PR 08-APR-1999; 99US-0128714P.

PR 16-APR-1999; 99US-0128845P.

PR 19-APR-1999; 99US-0130077P.

PR 21-APR-1999; 99US-0130449P.

PR 23-APR-1999; 99US-0130510P.

PR 23-APR-1999; 99US-0130891P.

PR 28-APR-1999; 99US-0131443P.

PR 30-APR-1999; 99US-0132048P.

PR 30-APR-1999; 99US-0132407P.

PR 04-MAY-1999; 99US-0132484P.

PR 05-MAY-1999; 99US-0132485P.

PR 06-MAY-1999; 99US-0132486P.

PR 06-MAY-1999; 99US-0132487P.

PR 07-MAY-1999; 99US-0132863P.

PR 11-MAY-1999; 99US-0134256P.

PR 14-MAY-1999; 99US-0134218P.

PR 14-MAY-1999; 99US-0134219P.

PR 14-MAY-1999; 99US-0134221P.

PR 14-MAY-1999; 99US-0134370P.

PR 18-MAY-1999; 99US-0134768P.

PR 19-MAY-1999; 99US-0134941P.

PR 20-MAY-1999; 99US-0135124P.

PR 21-MAY-1999; 99US-0135353P.

PR 24-MAY-1999; 99US-0135629P.

PR 25-MAY-1999; 99US-0136021P.

PR 27-MAY-1999; 99US-0136392P.
 PR 28-MAY-1999; 99US-0136782P.
 PR 01-JUN-1999; 99US-0137222P.
 PR 03-JUN-1999; 99US-0137528P.
 PR 04-JUN-1999; 99US-0137502P.
 PR 07-JUN-1999; 99US-0137724P.
 PR 08-JUN-1999; 99US-0138094P.
 PR 10-JUN-1999; 99US-0138540P.
 PR 10-JUN-1999; 99US-0138847P.
 PR 14-JUN-1999; 99US-0139119P.
 PR 16-JUN-1999; 99US-0139452P.
 PR 16-JUN-1999; 99US-0139453P.
 PR 17-JUN-1999; 99US-0139492P.
 PR 18-JUN-1999; 99US-0139454P.
 PR 18-JUN-1999; 99US-0139455P.
 PR 18-JUN-1999; 99US-0139456P.
 PR 18-JUN-1999; 99US-0139457P.
 PR 18-JUN-1999; 99US-0139458P.
 PR 18-JUN-1999; 99US-0139459P.
 PR 18-JUN-1999; 99US-0139460P.
 PR 18-JUN-1999; 99US-0139461P.
 PR 18-JUN-1999; 99US-0139462P.
 PR 18-JUN-1999; 99US-0139463P.
 PR 18-JUN-1999; 99US-0139750P.
 PR 18-JUN-1999; 99US-0139763P.
 PR 21-JUN-1999; 99US-0139817P.
 PR 22-JUN-1999; 99US-0139899P.
 PR 23-JUN-1999; 99US-0140353P.
 PR 23-JUN-1999; 99US-0140354P.
 PR 24-JUN-1999; 99US-0140695P.
 PR 28-JUN-1999; 99US-0140823P.
 PR 29-JUN-1999; 99US-0140931P.
 PR 30-JUN-1999; 99US-0141287P.
 PR 01-JUL-1999; 99US-0141842P.
 PR 02-JUL-1999; 99US-0142154P.
 PR 06-JUL-1999; 99US-0142055P.
 PR 08-JUL-1999; 99US-0142803P.
 PR 09-JUL-1999; 99US-0142920P.
 PR 12-JUL-1999; 99US-0142977P.
 PR 13-JUL-1999; 99US-0143542P.
 PR 14-JUL-1999; 99US-0143624P.
 PR 15-JUL-1999; 99US-0144005P.
 PR 16-JUL-1999; 99US-0144085P.
 PR 16-JUL-1999; 99US-0144086P.
 PR 19-JUL-1999; 99US-0144325P.
 PR 19-JUL-1999; 99US-0144331P.
 PR 19-JUL-1999; 99US-0144332P.
 PR 19-JUL-1999; 99US-0144333P.
 PR 19-JUL-1999; 99US-0144334P.
 PR 19-JUL-1999; 99US-0144335P.
 PR 20-JUL-1999; 99US-0144352P.
 PR 20-JUL-1999; 99US-0144632P.
 PR 20-JUL-1999; 99US-0144884P.
 PR 21-JUL-1999; 99US-0144814P.
 PR 21-JUL-1999; 99US-0145086P.
 PR 21-JUL-1999; 99US-0145088P.
 PR 22-JUL-1999; 99US-0145085P.
 PR 22-JUL-1999; 99US-0145087P.
 PR 22-JUL-1999; 99US-0145089P.
 PR 22-JUL-1999; 99US-0145192P.
 PR 23-JUL-1999; 99US-0145145P.
 PR 23-JUL-1999; 99US-0145218P.
 PR 23-JUL-1999; 99US-0145224P.
 PR 26-JUL-1999; 99US-0145276P.
 PR 27-JUL-1999; 99US-0145913P.
 PR 27-JUL-1999; 99US-0145918P.
 PR 27-JUL-1999; 99US-0145919P.
 PR 28-JUL-1999; 99US-0145951P.
 PR 02-AUG-1999; 99US-0146386P.
 PR 02-AUG-1999; 99US-0146388P.
 PR 02-AUG-1999; 99US-0146389P.
 PR 03-AUG-1999; 99US-0147038P.

```

PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.

PR 29-OCT-1999; 99US-0162142P.

Query Match 100.0%; Score 45; DB 3; Length 322;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
Db 87 RPQPPPP 93

RESULT 15
AAG06371
ID AAG06371 standard; protein; 340 AA.
XX
AC AAG06371;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 3120.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; Genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 25-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.

```

PR 10-JUN-1999; 99US-0138540P.
 PR 10-JUN-1999; 99US-0138847P.
 PR 14-JUN-1999; 99US-0139111P.
 PR 16-JUN-1999; 99US-0139452P.
 PR 16-JUN-1999; 99US-0139453P.
 PR 17-JUN-1999; 99US-0139492P.
 PR 18-JUN-1999; 99US-0139454P.
 PR 18-JUN-1999; 99US-0139455P.
 PR 18-JUN-1999; 99US-0139456P.
 PR 18-JUN-1999; 99US-0139457P.
 PR 18-JUN-1999; 99US-0139458P.
 PR 18-JUN-1999; 99US-0139459P.
 PR 18-JUN-1999; 99US-0139460P.
 PR 18-JUN-1999; 99US-0139461P.
 PR 18-JUN-1999; 99US-0139462P.
 PR 18-JUN-1999; 99US-0139463P.
 PR 18-JUN-1999; 99US-0139750P.
 PR 18-JUN-1999; 99US-0139763P.
 PR 21-JUN-1999; 99US-0139817P.
 PR 22-JUN-1999; 99US-0139899P.
 PR 23-JUN-1999; 99US-0140353P.
 PR 23-JUN-1999; 99US-0140354P.
 PR 24-JUN-1999; 99US-0140695P.
 PR 28-JUN-1999; 99US-0140823P.
 PR 29-JUN-1999; 99US-0140991P.
 PR 30-JUN-1999; 99US-0141287P.
 PR 01-JUL-1999; 99US-0141842P.
 PR 01-JUL-1999; 99US-0142154P.
 PR 02-JUL-1999; 99US-0142055P.
 PR 06-JUL-1999; 99US-0142390P.
 PR 08-JUL-1999; 99US-0142803P.
 PR 09-JUL-1999; 99US-0142920P.
 PR 12-JUL-1999; 99US-0142977P.
 PR 13-JUL-1999; 99US-0143542P.
 PR 14-JUL-1999; 99US-0143624P.
 PR 15-JUL-1999; 99US-0144005P.
 PR 16-JUL-1999; 99US-0144085P.
 PR 16-JUL-1999; 99US-0144086P.
 PR 19-JUL-1999; 99US-0144325P.
 PR 19-JUL-1999; 99US-0144331P.
 PR 19-JUL-1999; 99US-0144332P.
 PR 19-JUL-1999; 99US-0144333P.
 PR 19-JUL-1999; 99US-0144334P.
 PR 19-JUL-1999; 99US-0144335P.
 PR 20-JUL-1999; 99US-0144352P.
 PR 20-JUL-1999; 99US-0144632P.
 PR 21-JUL-1999; 99US-0144884P.
 PR 21-JUL-1999; 99US-0145086P.
 PR 21-JUL-1999; 99US-0145088P.
 PR 22-JUL-1999; 99US-0145085P.
 PR 22-JUL-1999; 99US-0145087P.
 PR 22-JUL-1999; 99US-0145089P.
 PR 22-JUL-1999; 99US-0145192P.
 PR 23-JUL-1999; 99US-0145145P.
 PR 23-JUL-1999; 99US-01452218P.
 PR 23-JUL-1999; 99US-0145224P.
 PR 26-JUL-1999; 99US-0145276P.
 PR 27-JUL-1999; 99US-0145913P.
 PR 27-JUL-1999; 99US-0145918P.
 PR 27-JUL-1999; 99US-0145919P.
 PR 28-JUL-1999; 99US-0145951P.
 PR 02-AUG-1999; 99US-0146386P.
 PR 02-AUG-1999; 99US-0146388P.
 PR 02-AUG-1999; 99US-0146389P.
 PR 03-AUG-1999; 99US-0147038P.
 PR 04-AUG-1999; 99US-0147204P.
 PR 04-AUG-1999; 99US-0147302P.
 PR 05-AUG-1999; 99US-0147260P.
 PR 06-AUG-1999; 99US-0147303P.
 PR 06-AUG-1999; 99US-0147416P.
 PR 09-AUG-1999; 99US-0147493P.

PR 09-AUG-1999; 99US-0147935P.
 PR 10-AUG-1999; 99US-0148171P.
 PR 11-AUG-1999; 99US-0148319P.
 PR 12-AUG-1999; 99US-0148341P.
 PR 13-AUG-1999; 99US-0148565P.
 PR 13-AUG-1999; 99US-0148684P.
 PR 16-AUG-1999; 99US-0149388P.
 PR 17-AUG-1999; 99US-0149175P.
 PR 18-AUG-1999; 99US-0149426P.
 PR 20-AUG-1999; 99US-0149722P.
 PR 20-AUG-1999; 99US-0149723P.
 PR 20-AUG-1999; 99US-0149929P.
 PR 23-AUG-1999; 99US-0149902P.
 PR 23-AUG-1999; 99US-0149930P.
 PR 25-AUG-1999; 99US-0150566P.
 PR 26-AUG-1999; 99US-0150884P.
 PR 27-AUG-1999; 99US-0151065P.
 PR 27-AUG-1999; 99US-0151066P.
 PR 27-AUG-1999; 99US-0151080P.
 PR 30-AUG-1999; 99US-0151303P.
 PR 31-AUG-1999; 99US-0151438P.
 PR 01-SEP-1999; 99US-0151930P.
 PR 07-SEP-1999; 99US-0152363P.
 PR 10-SEP-1999; 99US-0153070P.
 PR 13-SEP-1999; 99US-0153758P.
 PR 15-SEP-1999; 99US-0154018P.
 PR 16-SEP-1999; 99US-0154039P.
 PR 20-SEP-1999; 99US-0154779P.
 PR 22-SEP-1999; 99US-0155139P.
 PR 23-SEP-1999; 99US-0155486P.
 PR 24-SEP-1999; 99US-0155659P.
 PR 28-SEP-1999; 99US-0156458P.
 PR 29-SEP-1999; 99US-0156596P.
 PR 04-OCT-1999; 99US-0157117P.
 PR 05-OCT-1999; 99US-0157753P.
 PR 06-OCT-1999; 99US-0157865P.
 PR 07-OCT-1999; 99US-0158029P.
 PR 08-OCT-1999; 99US-0158232P.
 PR 12-OCT-1999; 99US-0158369P.
 PR 13-OCT-1999; 99US-0159293P.
 PR 13-OCT-1999; 99US-0159294P.
 PR 13-OCT-1999; 99US-0159295P.
 PR 14-OCT-1999; 99US-0159329P.
 PR 14-OCT-1999; 99US-0159330P.
 PR 14-OCT-1999; 99US-0159331P.
 PR 14-OCT-1999; 99US-0159637P.
 PR 14-OCT-1999; 99US-0159638P.
 PR 18-OCT-1999; 99US-0159584P.
 PR 21-OCT-1999; 99US-0160741P.
 PR 21-OCT-1999; 99US-0160767P.
 PR 21-OCT-1999; 99US-0160768P.
 PR 21-OCT-1999; 99US-0160770P.
 PR 21-OCT-1999; 99US-0160814P.
 PR 21-OCT-1999; 99US-0160815P.
 PR 22-OCT-1999; 99US-0160980P.
 PR 22-OCT-1999; 99US-0160981P.
 PR 22-OCT-1999; 99US-0160989P.
 PR 25-OCT-1999; 99US-0161404P.
 PR 25-OCT-1999; 99US-0161405P.
 PR 25-OCT-1999; 99US-0161406P.
 PR 26-OCT-1999; 99US-0161359P.
 PR 26-OCT-1999; 99US-0161360P.
 PR 28-OCT-1999; 99US-0161361P.
 PR 28-OCT-1999; 99US-0161920P.
 PR 28-OCT-1999; 99US-0161992P.
 PR 28-OCT-1999; 99US-0161993P.
 PR 29-OCT-1999; 99US-0162142P.

Query Match 100.0%; Score 45; DB 3; Length 340;
 Best Local Similarity 100.0%; Pred. No. 2.6e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 RQPPPP 7


```
Db      105 RQPPPP 111
|||||||
RESULT 16
AAG06370
ID AAG06370 standard; protein; 398 AA.
XX
AC AAG06370;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 3119.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
FN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-01213180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132407P.
PR 05-MAY-1999; 99US-0132484P.
PR 06-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 11-MAY-1999; 99US-0132863P.
PR 14-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
```

```
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 23-AUG-1999; 99US-0150566P.
PR 25-AUG-1999; 99US-0150884P.
PR 26-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153759P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156598P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157533P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160816P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 100.0%; Score 45; DB 3; Length 398;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
DB 163 RPQPPPP 169

RESULT 17
ADN73283
ID ADN73283 standard; protein; 398 AA.

XX 99US-0149175P.
AC 99US-0149426P.
XX 99US-0149722P.
DT 15-JUL-2004 (first entry)
XX 99US-0149723P.
DE 99US-0149929P.
XX 99US-0149902P.
XX 99US-0149930P.
KW 99US-0150566P.
KW 99US-0150884P.
KW 99US-0151065P.
XX 99US-0151066P.
OS 99US-0151080P.
XX 99US-0151303P.
PN 99US-0151438P.
XX WO2004035798-A2.
XX 99US-0151930P.
PD 99US-0152363P.
XX 99US-0153070P.
PF 99US-0153759P.
XX 99US-0154018P.
XX 99US-0154039P.
XX (CROP-) CROPDESIGN NV.
XX Inze D, De Veylder L, Vlieghe K;
PI WPT; 2004-348466/32.
DR N-PSDB; ADN73282.
XX
XX Altering plant characteristics, useful for producing plants for enzyme or
pharmaceutical production comprises modifying in a plant, expression of
one or more nucleic acids and/or modifying level or activity of one or
more proteins.
XX Claim 1; SEQ ID NO 1178; 134pp; English.
XX This invention relates to a novel method for altering one or more plant
characteristics. Specifically, it refers to identifying genes that are up
- or down-regulated in transgenic plants overexpressing the heterodimeric
E2Fa/Dpa transcription factor of Arabidopsis and using these sequences to
alter plant characteristics accordingly. The present invention describes
generating transgenic plants for the production of growth regulators,
enzymes, therapeutics, pharmaceuticals and animal feed products, where
the altered plant characteristics are selected from increased yield or
biomass, enhanced survival capacity, stress tolerance, plant architecture
or physiology, altered endoreplication, biochemistry, signal
transduction, storage lipid mobilisation and/or altered photosynthesis,
each relative to the corresponding wild type plants. Accordingly, these
sequences can also be useful as positive or negative selectable markers
during transformation of cells or tissues. The identified genes play a
role in a variety of biological processes such as DNA replication, cell
wall biosynthesis, nitrogen and/ or carbon metabolism or they function as
transcription factors. This polypeptide sequence is thale cress protein
expressed by a gene upregulated 1.3 fold or more in plants overexpressing
the E2Fa/Dpa transcription factor, given in an exemplification of the
invention.
XX Sequence 398 AA;
SQ Query Match 100.0%; Score 45; DB 8; Length 398;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
DB 163 RPQPPPP 169

RESULT 18
AAG06369
ID AAG06369 standard; protein; 412 AA.
XX
AC AAG06369;
XX
```

DT	17-OCT-2000	(first entry)	PR	18-JUN-1999;	99US-0139763P.
XX			PR	21-JUN-1999;	99US-0139817P.
DE	Arabidopsis thaliana	protein fragment SEQ ID NO: 3118.	PR	22-JUN-1999;	99US-0139899P.
XX			PR	23-JUN-1999;	99US-0140353P.
KW	Protein identification;	signal transduction pathway; metabolic pathway;	PR	23-JUN-1999;	99US-0140354P.
KW	hybridisation assay;	genetic mapping; gene expression control; promoter;	PR	24-JUN-1999;	99US-0140695P.
KW	termination sequence.		PR	28-JUN-1999;	99US-0140823P.
XX			PR	29-JUN-1999;	99US-0140991P.
OS	Arabidopsis thaliana.		PR	30-JUN-1999;	99US-0141287P.
XX			PR	01-JUL-1999;	99US-0141842P.
PN	EP1033405-A2.		PR	01-JUL-1999;	99US-0142154P.
XX			PR	02-JUL-1999;	99US-0142055P.
PD			PR	06-JUL-1999;	99US-0142390P.
XX			PR	08-JUL-1999;	99US-0142803P.
PF	06-SEP-2000.		PR	09-JUL-1999;	99US-0142920P.
XX			PR	12-JUL-1999;	99US-0142977P.
XX	25-FEB-2000;	2000EP-00301439.	PR	13-JUL-1999;	99US-0143542P.
PR	25-FEB-1999;	99US-0121825P.	PR	14-JUL-1999;	99US-0143624P.
PR	05-MAR-1999;	99US-0123180P.	PR	15-JUL-1999;	99US-0144005P.
PR	09-MAR-1999;	99US-0123548P.	PR	16-JUL-1999;	99US-0144085P.
PR	23-MAR-1999;	99US-0125788P.	PR	16-JUL-1999;	99US-0144086P.
PR	25-MAR-1999;	99US-0126264P.	PR	19-JUL-1999;	99US-0144325P.
PR	29-MAR-1999;	99US-0126785P.	PR	19-JUL-1999;	99US-0144331P.
PR	01-APR-1999;	99US-0127462P.	PR	19-JUL-1999;	99US-0144332P.
PR	06-APR-1999;	99US-0128234P.	PR	19-JUL-1999;	99US-0144333P.
PR	08-APR-1999;	99US-0128714P.	PR	19-JUL-1999;	99US-0144334P.
PR	16-APR-1999;	99US-0129845P.	PR	19-JUL-1999;	99US-0144335P.
PR	19-APR-1999;	99US-0130077P.	PR	20-JUL-1999;	99US-0144352P.
PR	21-APR-1999;	99US-0130449P.	PR	20-JUL-1999;	99US-0144353P.
PR	23-APR-1999;	99US-0130510P.	PR	20-JUL-1999;	99US-0144354P.
PR	23-APR-1999;	99US-0130891P.	PR	20-JUL-1999;	99US-0144884P.
PR	28-APR-1999;	99US-0131449P.	PR	21-JUL-1999;	99US-0144814P.
PR	30-APR-1999;	99US-0132048P.	PR	21-JUL-1999;	99US-0145086P.
PR	30-APR-1999;	99US-0132407P.	PR	21-JUL-1999;	99US-0145088P.
PR	04-MAY-1999;	99US-0132484P.	PR	22-JUL-1999;	99US-0145085P.
PR	04-MAY-1999;	99US-0132485P.	PR	22-JUL-1999;	99US-0145087P.
PR	05-MAY-1999;	99US-0132486P.	PR	22-JUL-1999;	99US-0145089P.
PR	06-MAY-1999;	99US-0132487P.	PR	22-JUL-1999;	99US-0145192P.
PR	06-MAY-1999;	99US-0132488P.	PR	23-JUL-1999;	99US-0145145P.
PR	07-MAY-1999;	99US-0132863P.	PR	23-JUL-1999;	99US-0145218P.
PR	11-MAY-1999;	99US-0134256P.	PR	23-JUL-1999;	99US-0145224P.
PR	14-MAY-1999;	99US-0134218P.	PR	23-JUL-1999;	99US-0145224P.
PR	14-MAY-1999;	99US-0134219P.	PR	26-JUL-1999;	99US-0145276P.
PR	14-MAY-1999;	99US-0134221P.	PR	27-JUL-1999;	99US-0145913P.
PR	14-MAY-1999;	99US-0134370P.	PR	27-JUL-1999;	99US-0145918P.
PR	18-MAY-1999;	99US-0134768P.	PR	27-JUL-1999;	99US-0145919P.
PR	19-MAY-1999;	99US-0134941P.	PR	28-JUL-1999;	99US-0145951P.
PR	20-MAY-1999;	99US-0135124P.	PR	02-AUG-1999;	99US-0146386P.
PR	21-MAY-1999;	99US-0135353P.	PR	02-AUG-1999;	99US-0146388P.
PR	24-MAY-1999;	99US-0135629P.	PR	02-AUG-1999;	99US-0146389P.
PR	25-MAY-1999;	99US-0136021P.	PR	03-AUG-1999;	99US-0147038P.
PR	27-MAY-1999;	99US-0136392P.	PR	04-AUG-1999;	99US-0147204P.
PR	28-MAY-1999;	99US-0136782P.	PR	04-AUG-1999;	99US-0147302P.
PR	01-JUN-1999;	99US-0137222P.	PR	05-AUG-1999;	99US-0147192P.
PR	03-JUN-1999;	99US-0137528P.	PR	05-AUG-1999;	99US-0147260P.
PR	04-JUN-1999;	99US-0137502P.	PR	06-AUG-1999;	99US-0147303P.
PR	07-JUN-1999;	99US-0137724P.	PR	06-AUG-1999;	99US-0147416P.
PR	08-JUN-1999;	99US-0138094P.	PR	09-AUG-1999;	99US-0147493P.
PR	10-JUN-1999;	99US-0138540P.	PR	09-AUG-1999;	99US-0147935P.
PR	10-JUN-1999;	99US-0138847P.	PR	10-AUG-1999;	99US-0148171P.
PR	14-JUN-1999;	99US-0139119P.	PR	11-AUG-1999;	99US-0148319P.
PR	16-JUN-1999;	99US-0139452P.	PR	12-AUG-1999;	99US-0148341P.
PR	16-JUN-1999;	99US-0139453P.	PR	13-AUG-1999;	99US-0148565P.
PR	17-JUN-1999;	99US-0139492P.	PR	13-AUG-1999;	99US-0148684P.
PR	18-JUN-1999;	99US-0139454P.	PR	16-AUG-1999;	99US-0149368P.
PR	18-JUN-1999;	99US-0139455P.	PR	17-AUG-1999;	99US-0149175P.
PR	18-JUN-1999;	99US-0139456P.	PR	18-AUG-1999;	99US-0149426P.
PR	18-JUN-1999;	99US-0139457P.	PR	20-AUG-1999;	99US-0149722P.
PR	18-JUN-1999;	99US-0139458P.	PR	20-AUG-1999;	99US-0149723P.
PR	18-JUN-1999;	99US-0139459P.	PR	20-AUG-1999;	99US-0149929P.
PR	18-JUN-1999;	99US-0139460P.	PR	23-AUG-1999;	99US-0149902P.
PR	18-JUN-1999;	99US-0139461P.	PR	23-AUG-1999;	99US-0149930P.
PR	18-JUN-1999;	99US-0139462P.	PR	25-AUG-1999;	99US-0150566P.
PR	18-JUN-1999;	99US-0139463P.	PR	26-AUG-1999;	99US-0150884P.
PR	18-JUN-1999;	99US-0139750P.	PR	27-AUG-1999;	99US-0151065P.

```
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154033P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.
```

```
Query Match 100.0%; Score 45; DB 3; Length 412;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy 1 RPQPPPP 7
Db 177 RPQPPPP 183
```

```
RESULT 19
AAW72182
ID AAW72182 standard; protein; 432 AA.
```

```
XX AAW72182;
XX
XX
XX 13-JAN-1999 (first entry)
DE HSV-2 strain SB5 Contig ID 15 ORF#15 protein.
XX
XX HSV-2 strain SB5; immunological response induction; therapy;
KW antiviral identification; viral protein inhibitor.
XX
```

```
OS Herpes simplex virus 2.
XX WO9820016-A1.
XX 14-MAY-1998.
XX 31-OCT-1997; 97WO-US020016.
XX 04-NOV-1996; 96US-0030279P.
XX 09-JUN-1997; 97US-0049018P.
XX (SMIK ) SMITHKLINE BEECHAM CORP.
XX Esser KM, Chan JY, Dabrowski-Amara CE, Delvecchio AM, Dillon SB;
XX Leary JJ;
XX WPI; 1998-286847/25.
XX N-PSDB; AAV62176.
XX Herpes simplex virus type-2 sequences - useful in, e.g. prevention and
XX treatment of infection or inducing immunological response in mammal.
XX Claim 10; Page 115; 748pp; English.
XX This sequence represents a Herpes simplex virus type-2 (HSV-2) protein
XX sequence of the invention. This sequence was isolated from a HSV-2 strain
XX SB5 (deposited as ATCC VR-2546) DNA fragment designated Contig ID 15.
XX Based on homology, this sequence is a probable DNA packaging protein. The
XX proteins can be used for the treatment or prevention of disease, to
XX induce an immunological response in a mammal or to identify inhibitors,
XX activators or novel antivirals. Antagonists of the proteins can be used
XX to inhibit a viral polypeptide. The DNA sequence or a vector containing
XX it can also be used to induce an immunological response in a mammal
XX
XX SQ Sequence 432 AA;
```

```
Query Match 100.0%; Score 45; DB 2; Length 432;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy 1 RPQPPPP 7
Db 377 RPQPPPP 383
```

```
RESULT 20
ABB70063
ID ABB70063 standard; protein; 446 AA.
```

```
XX ABB70063;
```

```
XX 26-MAR-2002 (first entry)
XX
XX Drosophila melanogaster polypeptide SEQ ID NO 36981.
XX
XX Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical.
XX
XX Drosophila melanogaster.
```

```
XX WO200171042-A2.
XX
XX 27-SEP-2001.
```

```
XX 23-MAR-2001; 2001WO-US009231.
XX
XX 23-MAR-2000; 2000US-0191637P.
```

```
XX 11-JUL-2000; 2000US-00614150.
XX
XX (PEKE ) PE CORP NY.
```

```
XX Venter JC, Adams M, Li PWD, Myers EW;
XX
```

DR WPI; 2001-656860/75.
 XX N-PSDB; ABL14166.
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signaling and cell-cell
 PT interactions.
 XX
 PS Disclosure; SEQ ID NO 36981; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
 CC ABB72072). The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 446 AA;
 Query Match 100.0%; Score 45; DB 4; Length 446;
 Best Local Similarity 100.0%; Pred. No. 3.2e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 RPQPPPP 7
 Db 231 RPQPPPP 237
 RESULT 21
 AAW72030
 ID AAW72030 standard; protein; 497 AA.
 XX AC AAW72030;
 XX
 DT 07-DEC-1998 (first entry)
 XX
 DE HSV-2 strain SB5 Contig ID 103 ORF#8 protein.
 XX
 KW HSV-2 strain SB5; immunological response induction; therapy;
 KW antiviral identification; viral protein inhibitor.
 XX
 OS Herpes simplex virus 2.
 XX
 PN WO9820016-A1.
 XX
 PD 14-MAY-1998.
 XX
 PF 31-OCT-1997; 97WO-US020016.
 XX
 PR 04-NOV-1996; 96US-0030279P.
 PR 09-JUN-1997; 97US-0049018P.
 XX
 PA (SMIK) SMITHKLINE BEECHAM CORP.
 XX
 PI Esser KM, Chan JY, Dabrowski-Amaral CE, Delvecchio AM, Dillon SB;
 PI Leary JJ;
 XX
 OS Herpes simplex virus 2.
 XX
 PN WO9820016-A1.
 XX
 PD 14-MAY-1998.
 XX
 PF 31-OCT-1997; 97WO-US020016.
 XX
 PR 04-NOV-1996; 96US-0030279P.
 PR 09-JUN-1997; 97US-0049018P.
 XX
 PA (SMIK) SMITHKLINE BEECHAM CORP.
 XX
 PI Esser KM, Chan JY, Dabrowski-Amaral CE, Delvecchio AM, Dillon SB;
 PI Leary JJ;
 XX
 DR WPI; 1998-286847/25.
 DR N-PSDB; AAV62133.
 XX
 PT Herpes simplex virus type-2 sequences - useful in, e.g. prevention and
 PT treatment of infection or inducing immunological response in mammal.
 XX
 PS Claim 10; Page 50; 748pp; English.
 XX
 CC This sequence represents a Herpes simplex virus type-2 (HSV-2) protein
 CC sequence of the invention. This sequence was isolated from a HSV-2 strain
 CC SB5 (deposited as ATCC VR-2546) DNA fragment designated Contig ID 103.
 CC Based on homology, this sequence is a 38K protein - human herpesvirus 1.
 CC The proteins can be used for the treatment or prevention of disease, to

CC induce an immunological response in a mammal or to identify inhibitors,
 CC activators or novel antivirals. Antagonists of the proteins can be used
 CC to inhibit a viral polypeptide. The DNA sequence or a vector containing
 CC it can also be used to induce an immunological response in a mammal
 XX
 SQ Sequence 497 AA;
 Query Match 100.0%; Score 45; DB 2; Length 497;
 Best Local Similarity 100.0%; Pred. No. 3.5e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 RPQPPPP 7
 Db 442 RPQPPPP 448
 RESULT 22
 AAW72132
 ID AAW72132 standard; protein; 497 AA.
 XX AC AAW72132;
 XX
 DT 23-DEC-1998 (first entry)
 XX
 DE HSV-2 strain SB5 Contig ID 16 ORF#8 protein.
 XX
 KW HSV-2 strain SB5; immunological response induction; therapy;
 KW antiviral identification; viral protein inhibitor.
 XX
 OS Herpes simplex virus 2.
 XX
 PN WO9820016-A1.
 XX
 PD 14-MAY-1998.
 XX
 PF 31-OCT-1997; 97WO-US020016.
 XX
 PR 04-NOV-1996; 96US-0030279P.
 PR 09-JUN-1997; 97US-0049018P.
 XX
 PA (SMIK) SMITHKLINE BEECHAM CORP.
 XX
 PI Esser KM, Chan JY, Dabrowski-Amaral CE, Delvecchio AM, Dillon SB;
 PI Leary JJ;
 XX
 DR WPI; 1998-286847/25.
 DR N-PSDB; AAV62160.
 XX
 PT Herpes simplex virus type-2 sequences - useful in, e.g. prevention and
 PT treatment of infection or inducing immunological response in mammal.
 XX
 PS Claim 10; Page 95; 748pp; English.
 XX
 CC This sequence represents a Herpes simplex virus type-2 (HSV-2) protein
 CC sequence of the invention. This sequence was isolated from a HSV-2 strain
 CC SB5 (deposited as ATCC VR-2546) DNA fragment designated Contig ID 16.
 CC Based on homology, this sequence is a probable DNA packaging protein. The
 CC proteins can be used for the treatment or prevention of disease, to
 CC induce an immunological response in a mammal or to identify inhibitors,
 CC activators or novel antivirals. Antagonists of the proteins can be used
 CC to inhibit a viral polypeptide. The DNA sequence or a vector containing
 CC it can also be used to induce an immunological response in a mammal
 XX
 SQ Sequence 497 AA;
 Query Match 100.0%; Score 45; DB 2; Length 497;
 Best Local Similarity 100.0%; Pred. No. 3.5e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 RPQPPPP 7
 Db 442 RPQPPPP 448

```

RESULT 23
AAZ29515
ID AAZ29515 standard; protein; 527 AA.
XX
XX AAZ29515;
AC
XX
XX 13-OCT-1999 (first entry)
DT
XX Human lung tumour protein SAL-50 2nd predicted amino acid sequence.
DE
XX Human; lung tumour protein; therapy; diagnosis; lung cancer; vaccine;
KW immunotherapy; detection; inhibition.
XX
XX Homo sapiens.
OS
XX WO9938973-A2.
PN
XX
XX 05-AUG-1999.
PD
XX
XX 26-JAN-1999; 99WO-US001642.
PF
XX
XX 28-JAN-1998; 98US-00015022.
PR
XX 28-JAN-1998; 98US-00015029.
PR
XX 18-MAR-1998; 98US-00040828.
PR
XX 18-MAR-1998; 98US-00040831.
PR
XX 23-JUL-1998; 98US-00122191.
PR
XX 23-JUL-1998; 98US-00122192.
PR
XX 22-DEC-1998; 98US-00219245.
XX
XX (CORI-) CORIXA CORP.
PA
XX
XX Reed SG, Lodes MJ, Frudakis TN, Mohamath R;
PI
XX WPI; 1999-479187/40.
DR
XX N-PSDB; AAZ07194.
DR
XX
XX Lung tumor specific polynucleotides for inhibiting the development of
PT lung cancer.
PT
XX
XX Claim 3; Page 170-171; 171pp; English.
PS
XX
XX The present invention describes lung tumour specific polynucleotides and
CC tumour antigens. AAZ07144 to AAZ07246 and AAZ08301 to AAZ08325 represent
CC specifically claimed polynucleotides, and AAZ29486 to AAZ29571 represent
CC amino acid sequences from the present invention. The lung tumour specific
CC polynucleotides and polypeptides can be used in pharmaceutical
CC polynucleotides and vaccines to inhibit the development of lung cancer. They
CC compositions and vaccines to detect lung cancer in a patient. Probes and
CC can also be used to detect lung cancer in a patient. Probes and
CC antibodies derived from the lung tumour sequences are useful in detection
CC of lung cancer
XX
XX Sequence 527 AA;
SQ
Query Match 100.0%; Score 45; DB 2; Length 527;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
DB 209 RPQPPPP 215

RESULT 24
AAB44492
ID AAB44492 standard; protein; 527 AA.
XX
XX AAB44492;
AC
XX
XX 05-FEB-2001 (first entry)
DT
XX Human lung tumour-specific antigen encoded by CDNA #100.
DE
XX

```

```

KW Lung tumour protein; lung cancer; cytostatic; vaccine.
XX
XX Homo sapiens.
OS
XX WO200060077-A2.
PN
XX
XX 12-OCT-2000.
PD
XX
XX 30-MAR-2000; 2000WO-US008560.
PF
XX
XX 02-APR-1999; 99US-00285323.
PR
XX 09-AUG-1999; 99US-00370838.
PR
XX 30-DEC-1999; 99US-00476235.
PR
XX 03-MAR-2000; 2000US-00518809.
XX
XX (CORI-) CORIXA CORP.
PA
XX
XX Reed SG, Lodes MJ, Mohamath R, Secrist H;
PI
XX
XX WPI; 2000-638466/61.
DR
XX N-PSDB; AAC79147.
DR
XX
XX Novel lung tumor polypeptides and polynucleotides, useful for detecting,
PT monitoring or treating cancer, especially lung cancer.
PT
XX
XX Example 4; Page 185-186; 243pp; English.
PS
XX
XX The present sequence is given in a specification relating to compounds
CC for therapy and diagnosis of lung cancer. Polypeptides comprising at
CC least an immunogenic part of a lung tumour protein are disclosed. The
CC polypeptides are useful for inhibiting the development of cancer,
CC especially lung cancer. Samples of T cells expressing the polypeptides
CC may be used to inhibit the development of cancer. The polypeptides are
CC also useful for detecting and monitoring the progression of cancer,
CC especially lung cancer
XX
XX Sequence 527 AA;
SQ
Query Match 100.0%; Score 45; DB 3; Length 527;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
DB 209 RPQPPPP 215

RESULT 25
AAE13805
ID AAE13805 standard; protein; 527 AA.
XX
XX AAE13805;
AC
XX
XX 26-FEB-2002 (first entry)
DT
XX
XX Human lung tumour-specific protein SAL-50 #2.
DE
XX
XX Human; lung tumour protein; immunostimulant; cytostatic; gene therapy;
KW antisense-therapy; vaccine; immune response; lung cancer; SAL-50.
XX
XX Homo sapiens.
OS
XX
XX WO200172295-A2.
PN
XX
XX 04-OCT-2001.
PD
XX
XX 28-MAR-2001; 2001WO-US009991.
PF
XX
XX 29-MAR-2000; 2000US-00538037.
PR
XX 05-JUN-2000; 2000US-00588937.
PR
XX 18-AUG-2000; 2000US-00640878.
PR
XX 22-SEP-2000; 2000US-0234517P.
PR
XX 01-NOV-2000; 2000US-00704512.

```

```

PR 14-DEC-2000; 2000US-00738973.
XX (CORI-) CORIXA CORP.
XX
XX Reed SG, Lodes MJ, Mohamath R, Secret H, Benson DR, Indirias CY;
XX Henderson RA, Fling SP, Algate PA, Elliot M, Mannion J, Kalos MD;
XX
XX WPI; 2001-639201/73.
XX N-PSDB; AAD23222.
XX
XX New human lung-specific polynucleotides and polypeptides for the
XX diagnosis and treatment of disease e.g. lung cancer.
XX
XX Example 4; Page 237-238; 378pp; English.
XX
XX The invention relates to isolated lung tumour-specific proteins and their
XX corresponding cDNA molecules. Lung tumour-specific proteins and their
XX antigen-presenting cells are useful for stimulating and/or expanding T
XX cells specific for a tumour protein, and for inhibiting the development
XX of cancer. The invention also relates to a composition useful for
XX stimulating an immune response, and for treating cancer. The lung tumour
XX specific oligonucleotide is useful in gene therapy and for diagnosis,
XX detection and treatment of lung cancer. The present sequence is human
XX lung tumour-specific protein
XX
XX SQ Sequence 527 AA;
XX
XX Query Match 100.0%; Score 45; DB 4; Length 527;
XX Best Local Similarity 100.0%; Pred. No. 3.7e+02;
XX Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 RPQPPPP 7
XX Db 209 RPQPPPP 215
XX
XX RESULT 26
XX ADD66524
XX ID ADD66524 standard; protein; 527 AA.
XX
XX AC ADD66524;
XX
XX DT 15-JAN-2004 (first entry)
XX
XX DE Human lung tumour-specific related protein, SEQ ID NO 216.
XX
XX KW expression control; cancer; T cell; tumour; immune; cytostatic; vaccine;
XX human; lung tumour-specific.
XX
XX OS Homo sapiens.
XX
XX PN WO200292001-A2.
XX
XX PD 21-NOV-2002.
XX
XX PF 10-MAY-2002; 2002WO-US014975.
XX
XX PR 11-MAY-2001; 2001US-00854133.
XX
XX PA (CORI-) CORIXA CORP.
XX
XX PI Lodes MJ, Wang T, Fan L, Algate PA, Mcneill PD;
XX WPI; 2003-120592/11.
XX
XX New polynucleotide and polypeptide, useful for preparing a composition
XX for diagnosing, treating or preventing cancer.
XX
XX Example 4; SEQ ID NO 216; 494pp; English.
XX
XX The invention relates to a novel isolated polynucleotide comprising one
XX of 32 47-6080 base pair sequences, given in the specification, or their
XX complements or degenerate variants, at least 20 contiguous residues of a

```

```

CC sequence in, or having at least 75 or 90 % identity with the isolated
CC polynucleotide, or that hybridise with the polynucleotide. The invention
CC further comprises: an isolated polypeptide; an expression vector
CC comprising the polynucleotide operably linked to an expression control
CC sequence; a host cell transformed or transfected with the expression
CC vector; an isolated antibody or its antigen-binding fragment that
CC specifically binds to the polypeptide; a method for detecting the
CC presence of a cancer in a patient; a fusion protein comprising the
CC polypeptide; an oligonucleotide that hybridises to the isolated
CC polynucleotide under moderately stringent conditions; a method for
CC stimulating and/or expanding T cells specific for a tumour protein; an
CC isolated T cell population; a composition comprising a first component
CC consisting of carriers and immunostimulants and a second component; a
CC method for stimulating an immune response in a patient; a method for
CC treating cancer in a patient; a method for determining cancer in a
CC patient; a diagnostic kit comprising at least one oligonucleotide or
CC antibody and a detection reagent comprising a reporter group; and a
CC method for inhibiting the development of cancer in a patient. The
CC compositions of the invention have cytostatic activity and can be used to
CC create a vaccine. The isolated polynucleotide is useful for preparing a
CC composition for diagnosing, treating or preventing cancer. This sequence
CC represents a human lung tumour-specific protein relating to the
CC invention.
XX
XX SQ Sequence 527 AA;
XX
XX Query Match 100.0%; Score 45; DB 7; Length 527;
XX Best Local Similarity 100.0%; Pred. No. 3.7e+02;
XX Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 RPQPPPP 7
XX Db 209 RPQPPPP 215
XX
XX RESULT 27
XX ADE87778
XX ID ADE87778 standard; protein; 527 AA.
XX
XX AC ADE87778;
XX
XX DT 29-JAN-2004 (first entry)
XX
XX DE Human lung tumour antigen polypeptide #86.
XX
XX KW Human; lung tumour antigen; cancer; lung cancer; CD4+; CD8+; T cell;
XX immune response; immunostimulant; cytostatic.
XX
XX OS Homo sapiens.
XX
XX PN US2003118599-A1.
XX
XX PD 26-JUN-2003.
XX
XX PF 10-MAY-2002; 2002US-00144649.
XX
XX PR 02-APR-1999; 99US-00285323.
XX PR 09-AUG-1999; 99US-00370838.
XX PR 30-DEC-1999; 99US-00476235.
XX PR 03-MAR-2000; 2000US-00518809.
XX PR 29-MAR-2000; 2000US-00538037.
XX PR 05-JUN-2000; 2000US-00588937.
XX PR 18-AUG-2000; 2000US-00640878.
XX PR 20-SEP-2000; 2000US-00667170.
XX PR 01-NOV-2000; 2000US-00704512.
XX PR 14-DEC-2000; 2000US-00738973.
XX PR 11-MAY-2001; 2001US-00854133.
XX
XX (CORI-) CORIXA CORP.
XX
XX Algate PA, Lodes MJ, Wang T, Fan L, Mcneill PD;
XX WPI; 2003-897103/82.
XX
XX

```

```

DR N-PSDB; ADB87743.
XX
XX New polynucleotides encode lung tumor antigens and are useful to
PT stimulate an immune response or detect or treat a cancer in a patient,
PT particularly lung cancer.
XX
XX Example 4; SEQ ID NO 216; 63pp; English.
XX
XX The invention relates to polynucleotides encoding lung tumour antigens.
XX The invention also relates to the polypeptides encoded by the
CC polynucleotides, isolated antibodies or antigen-binding fragments that
CC specifically bind the polypeptides and a method for detecting cancer in a
CC patient, comprising obtaining a biological sample from the patient,
CC contacting the sample with a binding agent that binds a polypeptide of
CC the invention, detecting in the sample an amount of polypeptide that
CC binds to the binding agent, and comparing the amount of polypeptide to a
CC predetermined cut-off value. T cells specific for a tumour protein can be
CC stimulated and/or expanded by contacting the T cells with a polypeptide,
CC polynucleotide or an antigen-presenting cell that expresses a
CC polypeptide. Cancer development can be inhibited by incubating CD4+
CC and/or CD8+ T cells isolated from a patient with a polypeptide,
CC polynucleotide or an antigen-presenting cell that expresses a
CC polypeptide, so that the T cells proliferate. The invention is used to
CC stimulate an immune response or to detect or treat a cancer in a patient,
CC particularly lung cancer. This sequence represents a human lung tumour
CC antigen polypeptide of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification but was obtained in
CC electronic format from USPTO at seqdata.uspto.gov/sequence.html.
XX
XX Sequence 527 AA;
SQ
Query Match 100.0%; Score 45; DB 7; Length 527;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RPQPPPP 7
DB 209 RPQPPPP 215
RESULT 28
AAU10545
ID AAU10545 standard; protein; 668 AA.
XX
XX AAU10545;
XX
XX 14-FEB-2002 (first entry)
XX
XX Rat synapsin 1B (YSG8) polypeptide.
XX
XX YSG; YSG8; schizophrenia; chronic animal model; LCGU; netrin receptor;
KW local cerebral glucose utilisation; phosphodiesterase 1-alpha; UNC5H1;
KW calcium-independent alpha-latrotoxin receptor; CIRL; trkE; synapsin 1A;
KW epithelial discoidin domain receptor 1; synapsin 1B; neuroleptic;
KW tumour necrosis factor alpha; TNF-alpha; rat.
XX
XX Rattus sp.
XX
XX WO200175440-A2.
XX
XX 11-OCT-2001.
XX
XX 02-APR-2001; 2001WO-GB001486.
XX
XX 31-MAR-2000; 2000GB-00007880.
XX
XX 26-MAY-2000; 2000GB-00012768.
XX
XX (WELF-) WELFIDE CORP.
XX
XX Cochran S, Paterson G, Ohashi Y, Morris B, Pratt J;
XX WPI; 2002-010813/01.
XX
XX N-PSDB; AAS16845.
XX
XX Novel chronic animal model of schizophrenia, useful for identifying anti-
PT psychotic drugs and genes that are associated with schizophrenia.
XX
XX Claim 1; Fig 9d; 79pp; English.
XX
XX The invention relates to YSG polynucleotide fragments for use in
CC diagnosing and/or developing treatments for schizophrenia using chronic
CC animal models. The polynucleotides and their encoded polypeptides are
CC used for identification of compounds which modulate the expression of YSG
CC molecules, leading to the manufacture of schizophrenia medications. The
CC sequences can also be used for testing candidate compounds for any effect
CC on the polypeptides. Anti-schizophrenic effects of a compound can be
CC determined by measuring local cerebral glucose utilisation (LCGU) or
CC comparing its expression level with that of a control group. The
CC sequences are useful in the identification of genes associated with
CC schizophrenic states and in the development of an antibody. The sequences
CC of the invention include phosphodiesterase 1-alpha, calcium-independent
CC alpha-latrotoxin receptors (CIRL)-1.2&3, epithelial discoidin domain
CC receptor 1 (trkE), netrin receptor (UNC5H1), synapsin 1A and AB and
CC tumour necrosis factor (TNF) alpha. This sequence represents rat synapsin
CC 1B (YSG8) polypeptide
XX
XX Sequence 668 AA;
SQ
Query Match 100.0%; Score 45; DB 5; Length 668;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RPQPPPP 7
DB 27 RPQPPPP 33
RESULT 29
ADX05987
ID ADX05987 standard; protein; 669 AA.
XX
XX ADX05987;
XX
XX 21-APR-2005 (first entry)
XX
XX Cyclin-dependent kinase modulation biomarker SEQ ID NO 552.
XX
XX cytosolic; cyclin-dependent kinase; cdk; biomarker.
XX
XX Homo sapiens.
XX
XX WO2005012875-A2.
XX
XX 10-FEB-2005.
XX
XX 29-JUL-2004; 2004WO-US024424.
XX
XX 29-JUL-2003; 2003US-0490890P.
XX
XX (BRIM ) BRISTOL-MYERS SQUIBB CO.
XX
XX Li M, Rupnow BA, Webster KR, Jackson DG, Wong TW;
XX WPI; 2005-163068/17.
XX
XX N-PSDB; ADX05986.
XX
XX Biomarkers useful for predicting or determining the response of a mammal
PT to a cancer treatment comprising administration of a modulator of cyclin-
PT dependent kinase activity.
XX
XX Claim 5; SEQ ID NO 552; 141pp; English.
XX
XX This invention describes a novel method of predicting or determining
CC whether a mammal will respond or is responding to an anti-cancer agent
CC that modulates cyclin-dependent kinase (cdk) activity. The method
CC comprises measuring the level of one or more biomarkers selected from

```


2774 biomarkers given in the specification (nucleotide sequence SEQ ID NO:1246 (Genbank EST W28729) is especially preferred). The method of the invention is utilized in a kit for determining or predicting whether patient would be susceptible or resistant to treatment by an agent modulating cdk activity. The invention also describes a method for utilizing individualized genetic profiles for treating diseases and disorders based on patient's response and molecular level, specialized microarrays comprising the biomarkers described, antibodies directed against the biomarkers and a cell culture model to identify biomarkers. The cdk modulator is preferably N-5-[(5-(1,1-Dimethylethyl)-2-oxazolyl)methyl]thiol-2-thiazolyl-4-piperidine carboxamide, 0.5-L-tartaric acid salt. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at fp.wipo.int/pub/published_pct_sequences. This sequence represents a biomarker used in the method of the invention.

Sequence 669 AA;
Query Match 100.0%; Score 45; DB 9; Length 669;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
Db 27 RPQPPPP 33

RESULT 30
AAU10544
ID AAU10544 standard; protein; 704 AA.

XX AC AAU10544;

DT 14-FEB-2002 (first entry)

DE Rat synapsin 1A (YSG8) polypeptide.

XX YSG; YSG8; schizophrenia; chronic animal model; LCGU; netrin receptor;
KW local cerebral glucose utilisation; phosphodiesterase 1-alpha; UNC5H1;
KW calcium-independent alpha-latrotoxin receptor; CIRL; trke; synapsin 1A;
KW epithelial discoidin domain receptor 1; synapsin 1B; neuroleptic;
KW tumour necrosis factor alpha; TNF-alpha; rat.

XX OS Rattus sp.

XX PN WO200175440-A2.

XX PD 11-OCT-2001.

XX PF 02-APR-2001; 2001WO-GB001486.

XX PR 31-MAR-2000; 2000GB-00007880.

XX PR 26-MAY-2000; 2000GB-00012768.

XX PA (WELF-) WELFIDE CORP.

XX PI Cochran S, Paterson G, Ohashi Y, Morris B, Pratt J;

XX DR WPI; 2002-010813/01.

XX DR N-PSDB; AAS16844.

XX Novel chronic animal model of schizophrenia, useful for identifying anti-psychotic drugs and genes that are associated with schizophrenia.

XX Disclosure; Fig 9b; 79pp; English.

XX The invention relates to YSG polynucleotide fragments for use in diagnosing and/or developing treatments for schizophrenia using chronic animal models. The polynucleotides and their encoded polypeptides are used for identification of compounds which modulate the expression of YSG molecules, leading to the manufacture of schizophrenia medicaments. The sequences can also be used for testing candidate compounds for any effect on the polypeptides. Anti-schizophrenic effects of a compound can be

CC determined by measuring local cerebral glucose utilisation (LCGU) or
CC comparing its expression level with that of a control group. The
CC sequences are useful in the identification of genes associated with
CC schizophrenic states and in the development of an antibody. The sequences
CC of the invention include phosphodiesterase 1-alpha, calcium-independent
CC alpha-latrotoxin receptors (CIRL)-1,2&3, epithelial discoidin domain
CC receptor 1 (trke), netrin receptor (UNC5H1), synapsins 1A and 1B and
CC tumour necrosis factor (TNF) alpha. This sequence represents rat synapsin
CC 1A (YSG8) polypeptide

XX Sequence 704 AA;

Query Match 100.0%; Score 45; DB 5; Length 704;
Best Local Similarity 100.0%; Pred. No. 4.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
Db 27 RPQPPPP 33

RESULT 31
ADE57712
ID ADE57712 standard; protein; 704 AA.

XX AC ADE57712;

DT 29-JAN-2004 (first entry)

DE Rat Protein P09951, SEQ ID NO 3575.

XX Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.

XX OS Rattus norvegicus.

XX PN WO2003016475-A2.

XX PD 27-FEB-2003.

XX PF 14-AUG-2002; 2002WO-US025765.

XX PR 14-AUG-2001; 2001US-0312147P.

XX PR 01-NOV-2001; 2001US-0346382P.

XX PR 26-NOV-2001; 2001US-0333347P.

XX PA (GEHO) GEN HOSPITAL CORP.

XX PA (FARB) BAYER AG.

XX PI Woolf C, D'urso D, Befort K, Costigan M;

XX WPI; 2003-268312/26.

XX DR GENBANK; P09951.

XX New composition comprising two or more isolated polypeptides, useful for preparing a medicament for treating pain in an animal.

XX Claim 1; Page; 101pp; English.

XX The invention discloses a composition comprising two or more isolated rat or human polynucleotides or a polynucleotide which represents a fragment, derivative or allelic variation of the nucleic acid sequence. Also claimed are a vector comprising the novel polynucleotide, a host cell comprising the vector, a method for identifying a nucleotide sequence which is differentially regulated in an animal subjected to pain and a kit to perform the method, an array, a method for identifying an agent that increases or decreases the expression of the polynucleotide sequence that is differentially expressed in neuronal tissue of a first animal subjected to pain, a method for identifying a compound which regulates the expression of a polynucleotide sequence which is differentially expressed in an animal subjected to pain, a method for identifying a compound that regulates the activity of one or more of the polynucleotides, a method for producing a pharmaceutical composition, a

CC method for identifying a compound or small molecule that regulates the
 CC activity in an animal of one or more of the polypeptides given in the
 CC specification, a method for identifying a compound useful in treating
 CC pain and a pharmaceutical composition comprising the one or more
 CC polypeptides or their antibodies. The polynucleotide or the compound that
 CC modulates its activity is useful for preparing a medicament for treating
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
 CC therapy). The sequence presented is a rat protein (shown in Table 2 of
 CC the specification) which is differentially expressed during pain. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic form directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 704 AA;
 SQ Query Match 100.0%; Score 45; DB 7; Length 704;
 Best Local Similarity 100.0%; Pred. No. 4.8e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPQPPPP 7
 |||||
 Db 27 RPQPPPP 33

RESULT 32

AD AM39328
 ID AM39328 standard; protein; 705 AA.

XX AM39328;

XX 22-OCT-2001 (first entry)

XX Human polypeptide SEQ ID NO 2473.

XX Human; neurotrophic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia.

XX Homo sapiens.

XX WO200153312-A1.

XX 26-JUL-2001.

XX 26-DEC-2000; 2000WO-US034263.

XX 23-DEC-1999; 99US-00471275.

XX 21-JAN-2000; 2000US-00488725.

XX 25-APR-2000; 2000US-0052317.

XX 20-JUN-2000; 2000US-00598042.

XX 19-JUL-2000; 2000US-00620312.

XX 03-AUG-2000; 2000US-00653450.

XX 14-SEP-2000; 2000US-00662191.

XX 19-OCT-2000; 2000US-00693036.

XX 29-NOV-2000; 2000US-00727344.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
 PI Zhou P, Goodrich R, Dermanac RT;
 XX WPI; 2001-442253/47.
 DR N-PSDB; AA158484.

XX Novel nucleic acids and polypeptides, useful for treating disorders such
 PT as central nervous system injuries.

XX Example 4; SEQ ID NO 2473; 10078pp; English.

XX The invention relates to human nucleic acids (AA157798-AA161369) and the
 CC encoded polypeptides (AA38642-AA42213) with neurotrophic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders. Note: The sequence data for this patent did not form
 CC part of the printed specification

XX Sequence 705 AA;

Query Match 100.0%; Score 45; DB 4; Length 705;

Best Local Similarity 100.0%; Pred. No. 4.8e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPQPPPP 7
 |||||

Db 27 RPQPPPP 33

RESULT 33

AD AM5985
 ID ADX05985 standard; protein; 705 AA.

XX ADX05985;

XX 21-APR-2005 (first entry)

XX Cyclin-dependent kinase modulation biomarker SEQ ID NO 550.

XX cytostatic; cyclin-dependent kinase; cdk; biomarker.

XX Homo sapiens.

XX WO2005012875-A2.

XX 10-FEB-2005.

XX 29-JUL-2004; 2004WO-US024424.

XX 29-JUL-2003; 2003US-0490890P.

XX (BRIM) BRISTOL-MYERS SQUIBB CO.

XX Li M, Rupnow BA, Webster KR, Jackson DG, Wong TW;

XX WPI; 2005-163068/17.

XX N-PSDB; ADX05984.

XX Biomarkers useful for predicting or determining the response of a mammal
 PT to a cancer treatment comprising administration of a modulator of cyclin-
 PT dependent kinase activity.

XX Claim 5; SEQ ID NO 550; 141pp; English.

XX This invention describes a novel method of predicting or determining
 CC whether a mammal will respond or is responding to an anti-cancer agent
 CC that modulates cyclin-dependent kinase (cdk) activity. The method
 CC comprises measuring the level of one or more biomarkers selected from
 CC 2774 biomarkers given in the specification (nucleotide sequence SEQ ID
 CC NO:1246 (Genbank EST W28729) is especially preferred). The method of the
 CC invention is utilized in a kit for determining or predicting whether
 CC patient would be susceptible or resistant to treatment by an agent
 CC modulating cdk activity. The invention also describes a method for
 CC utilizing individualized genetic profiles for treating diseases and

CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 735 AA;

Query Match 100.0%; Score 45; DB 4; Length 735;
 Best Local Similarity 100.0%; Pred. No. 4.9e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPQPPPP 7
 |||||
 Db 9 RPQPPPP 15

RESULT 36
 ABG30148
 ID ABG30148 standard; protein; 812 AA.

XX AC ABG30148;

XX DT 18-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #30139.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US008631.

XX PR 31-MAR-2000; 2000US-00540217.

XX PR 23-AUG-2000; 2000US-00649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX DR WPI; 2001-639362/73.

XX DR N-PSDB; AAS94335.

XX PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.

XX PS Claim 20; SEQ ID NO 60507; 103pp; English.

XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (II) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in

CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 812 AA;

Query Match 100.0%; Score 45; DB 4; Length 812;
 Best Local Similarity 100.0%; Pred. No. 5.4e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPQPPPP 7
 |||||
 Db 445 RPQPPPP 451

RESULT 37
 ABG2079
 ID ABG2079 standard; protein; 874 AA.

XX AC ABG2079;

XX DT 13-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #2070.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US008631.

XX PR 31-MAR-2000; 2000US-00540217.

XX PR 23-AUG-2000; 2000US-00649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX DR WPI; 2001-639362/73.

XX DR N-PSDB; AAS66266.

XX PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.

XX PS Claim 20; SEQ ID NO 32438; 103pp; English.

XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (II) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity

CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
 CC amino acid sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 XX Sequence 874 AA;
 SQ

Query Match 100.0%; Score 45; DB 4; Length 874;
 Best Local Similarity 100.0%; Pred. No. 5.7e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPQPPPP 7
 |||||
 Db 445 RPQPPPP 451

RESULT 38
 ABG00401
 ID ABG00401 standard; protein; 1074 AA.
 XX
 AC ABG00401;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #392.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO200175067-A2.
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US008631.
 XX
 PR 31-MAR-2000; 2000US-00540217.
 PR 23-AUG-2000; 2000US-00649167.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Drmanac RT, Liu C, Tang YT;
 XX
 DR WPI; 2001-639362/73.
 DR N-PSDB; AAS64588.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 XX
 PS Claim 20; SEQ ID NO 30760; 103pp; English.
 XX

The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic

CC amino acid sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 XX Sequence 1074 AA;
 SQ

Query Match 100.0%; Score 45; DB 4; Length 1074;
 Best Local Similarity 100.0%; Pred. No. 6.8e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPQPPPP 7
 |||||
 Db 445 RPQPPPP 451

RESULT 39
 AAE21729
 ID AAE21729 standard; protein; 1413 AA.
 XX
 AC AAE21729;
 XX
 DT 16-JUL-2002 (first entry)
 XX
 DE Human PKIN-24 protein.
 XX
 KW Human; kinase; enzyme; PKIN-24 protein; immune system disorder; anaemia;
 KW acquired immune deficiency syndrome; thymic hypoplasia; Crohn's disease;
 KW asthma; neurological disorder; epilepsy; Charcot-Marie-Tooth disease;
 KW AIDS; seizures; cell proliferative disorder; cancer; adenocarcinoma;
 KW leukaemia; lymphoma; melanoma; myeloma; sarcoma; developmental disorder;
 KW Down's syndrome; gene therapy; protein therapy; cytostatic.
 XX
 OS Homo sapiens.
 XX
 FH Key
 FT Peptide
 FT /label= Signal_peptide
 FT 21..1413
 FT /label= Human_mature_PKIN-19_protein
 FT 133..404
 FT /note= "Eukaryotic protein kinase domain"
 FT 136..400
 FT /note= "Protein kinase domain"
 FT 137..400
 FT /note= "Protein kinase domain"
 FT 137..398
 FT /note= "Protein kinase domain"
 FT 210..223
 FT /note= "Tyrosine kinase catalytic domain"
 FT 256..274
 FT /note= "Tyrosine kinase catalytic domain"
 FT 305..315
 FT /note= "Tyrosine kinase catalytic domain"
 FT 331..353
 FT /note= "Tyrosine kinase catalytic domain"
 FT 380..402
 FT /note= "Tyrosine kinase catalytic domain"
 PN WO200218557-A2.
 PD 07-MAR-2002.
 XX
 PF 31-AUG-2001; 2001WO-US027219.
 XX
 PR 31-AUG-2000; 2000US-0229873P.
 PR 08-SEP-2000; 2000US-0231357P.
 PR 14-SEP-2000; 2000US-0232654P.
 PR 22-SEP-2000; 2000US-0234902P.
 PR 29-SEP-2000; 2000US-0236499P.
 PR 06-OCT-2000; 2000US-0238389P.
 PR 13-OCT-2000; 2000US-0240542P.
 XX

```

PA (INCY-) INCYTE GENOMICS INC.
XX
PI Bandman O, Nguyen DB, Hafalia AJA, Yao MG, Gandhi AR,
PI Gururajan R, Ding L, Patterson C, Yue H, Baughn MR, Tribouley CM;
PI Thornton M, Elliott VS, Lu Y, Ison CH, Au-Young J, Tang YT;
PI Azimzai Y, Burdill JD, Marcus GA, Zingler KA, Lu DAM, Lal PG;
PI Rankumar J, Warren BA, Kearney L, Policky JL, Thangavelu K;
PI Burford N;
XX
DR WPI; 2002-329769/36.
DR N-PSDB; AAD34321.
XX
XX New human kinases, useful for diagnosing, treating or preventing immune
PT system disorders (e.g. Crohn's disease), neurological disorders (e.g.
PT epilepsy), or cell proliferative disorders (e.g. cancers such as leukemia
PT or lymphoma).
XX
PS Claim 79; Page 189-193; 218pp; English.
XX
XX The present invention relates to human kinases (PKIN) and polynucleotides
CC encoding such proteins. PKIN sequences of the invention are useful for
CC diagnosing, treating or preventing disorders associated with aberrant
CC expression of PKIN, particularly immune system disorders (e.g. acquired
CC immune deficiency syndrome (AIDS), thymic hypoplasia, Crohn's disease,
CC anaemia, asthma), neurological disorders (e.g. epilepsy, Charcot-Marie-
CC Tooth disease or seizures), cell proliferative disorders (e.g. cancers
CC such as adenocarcinoma, leukaemia, lymphoma, melanoma, sarcoma),
CC and developmental disorders (e.g. Down's syndrome). They are also used in
CC gene therapy and protein therapy. The present sequence is human PKIN-24
CC protein
XX
XX Sequence 1413 AA;
SQ
Query Match 100.0%; Score 45; DB 5; Length 1413;
Best Local Similarity 100.0%; Pred. No. 8.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 RPQPPPP 7
Db 1095 RPQPPPP 1101
|||||
|
RESULT 40
ADCS9465
ID ADC59465 standard; protein; 1460 AA.
XX
AC ADC59465;
XX
XX 18-DEC-2003 (first entry)
XX
DE Novel human tyrosine kinase protein.
XX
XX Enzyme; Tyrosine kinase; antiinflammatory; cytostatic;
KW antiarteriosclerotic; immunomodulatory; neuroprotective; gene therapy;
KW cell proliferation; Human.
XX
OS Homo sapiens.
XX
XX JP2003024075-A.
PN
XX
XX 28-JAN-2003.
PD
XX
XX 11-JUL-2001; 2001JP-00211135.
PF
XX
XX 11-JUL-2001; 2001JP-00211135.
PR
XX
XX (KAZU-) ZH KAZUSA DNA KENKYUSHO.
PA
XX
XX WPI; 2003-545691/52.
DR N-PSDB; ADC59464.
XX
XX Novel recombinant protein useful for controlling proliferation of various
PT cells, has biological activity substantially homogeneous as tyrosine
PT

```

```

PT kinase.
XX
PS Disclosure; SEQ ID NO 1; 22pp; Japanese.
XX
XX This invention relates to a novel recombinant protein which has tyrosine
CC kinase activity. The invention also discloses a pharmaceutical
CC composition comprising a DNA and protein sequence of the invention and an
CC antibody which binds specifically to the tyrosine kinase protein. The
CC protein of the invention may have antiinflammatory, cytostatic,
CC antiarteriosclerotic, immunomodulatory and neuroprotective activity and
CC may be used in gene therapy. The sequences of the invention are useful
CC for controlling proliferation of various cells, differentiation and
CC signal transduction. The protein is useful as a reagent for screening the
CC compound or its salt which inhibits the biological activity of the
CC tyrosine kinase. The sequences of the invention are also useful in the
CC diagnosis, prevention and treatment of diseases such as immune disorders,
CC neurological diseases, inflammation, cancer and arteriosclerosis. The
CC nucleic acid sequence of the invention may also useful as a detection
CC probe. The present sequence represents the novel human tyrosine kinase
CC protein of the invention.
XX
SQ Sequence 1460 AA;
Query Match 100.0%; Score 45; DB 7; Length 1460;
Best Local Similarity 100.0%; Pred. No. 8.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 RPQPPPP 7
Db 1142 RPQPPPP 1148
|||||
|
RESULT 41
ADU04632
ID ADU04632 standard; protein; 1460 AA.
XX
AC ADU04632;
XX
XX 13-JAN-2005 (first entry)
DT
XX
DE Human KIAA 1883 polypeptide.
XX
XX Human; KIAA 1883; lung cancer; marker; diagnosis.
XX
OS Homo sapiens.
XX
XX WO2004091511-A2.
PN
XX
XX 28-OCT-2004.
PD
XX
XX 12-APR-2004; 2004WO-US011193.
PF
XX
XX 10-APR-2003; 2003US-0462028P.
PR
XX
XX (GENZ ) GENZYME CORP.
PA
XX
XX Roberts BL;
PI
XX
XX WPI; 2004-766692/75.
DR N-PSDB; ADU04631.
XX
XX Detecting neoplasia in lung cells comprises detecting the level of
PT expression of at least one gene selected from EGFR-RS, RYK, TNFRS25,
PT TRPM7, UNC5H2, KCP3 and KIAA 1883.
XX
XX Disclosure; SEQ ID NO 14; 80pp; English.
PS
XX
XX The present sequence is that of human KIAA 1833 polypeptide. KIAA 1833
CC cDNA was identified in a sequencing project for large proteins of
CC unidentified genes. The invention provides a method of diagnosing the
CC neoplastic state of a lung cell by screening for the presence of a
CC differentially expressed gene. The gene is selected from EGFR-RS, RYK,
CC TNFRS25, TRPM7, KCP3 and KIAA 1833, which are expressed at higher levels
CC

```

CC in neoplastic lung cell or lung tumour cells than in healthy lung cells,
 CC or UNC5H2, which is expressed at a higher level in healthy lung cells
 CC than in neoplastic lung cells. The genes and their products provide
 CC diagnostic and prognostic markers and therapeutic targets. Databases
 CC containing quantitative full or partial transcripts or protein sequences
 CC isolated from a cell sample can be searched and analysed for the presence
 CC and amount of transcript or expressed gene product. The methods are
 CC particularly useful for aiding in the diagnosis of non-small cell lung
 CC cancer. The invention also provides methods of screening for potential
 CC therapeutic agents for the reversal of the neoplastic condition. Claimed
 CC methods for inhibiting the growth of a neoplastic lung cell comprise
 CC contacting the cell with a labelled antibody that specifically recognises
 CC and binds the protein product of at least one of the EGFR-RS, RYK,
 CC TNFRSF25, TRPM7, KCP3 or KIAA 1883 genes, or with an immune effector cell
 CC that specifically recognises and lyses a cell expressing at least one of
 CC the EGFR-RS, RYK, TNFRSF25, TRPM7, UNC5H2, KCP3 and KIAA 1883 genes.

XX
 SQ Sequence 1460 AA;
 Query Match 100.0%; Score 45; DB 8; Length 1460;
 Best Local Similarity 100.0%; Pred. No. 8.9e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
 |||||
 Db 1142 RPQPPPP 1148

RESULT 42
 ABR42259
 ID ABR42259 standard; protein; 1734 AA.
 XX
 AC ABR42259;
 XX
 DT 28-JUL-2003 (first entry)
 XX
 DE Human MAST205 protein.
 XX
 KW Human; MAST205; protein kinase; enzyme; cytostatic; antidiabetic;
 KW nootropic; neuroprotective; antiinflammatory; analgesic; gene therapy.
 XX
 OS Homo sapiens.

XX WO2003033709-A2.
 XX
 PD 24-APR-2003.
 XX
 PF 16-OCT-2002; 2002WO-EP011551.
 XX
 PR 17-OCT-2001; 2001US-0329502P.
 XX
 PA (FARB) BAYER AG.
 XX
 PI Zhu Z;
 XX
 DR WPI; 2003-393534/37.
 XX
 PT New polynucleotide encoding a serine/threonine protein kinase
 PT polypeptide, useful for preventing, ameliorating or treating diseases
 PT associated with the protein kinase dysfunction, e.g. cancer, diabetes or
 PT a CNS disorder.

XX
 PS Disclosure; Fig 6; 149pp; English.
 XX
 CC The present sequence is that of human MAST205 protein. A novel human
 CC serine/threonine protein kinase (see ABR42259) of the invention shows
 CC homology to this sequence. The serine/threonine protein kinase can be
 CC produced by recombinant methods and used to screen for reagents that
 CC modulate its activity. Such reagents are used in the preparation of a
 CC medicament for modulating the activity of the enzyme in a disease
 CC selected from cancer, diabetes, central nervous system disorders and
 CC chronic obstructive pulmonary disease (COPD) (claimed)

XX

SQ Sequence 1734 AA;

Query Match 100.0%; Score 45; DB 6; Length 1734;
 Best Local Similarity 100.0%; Pred. No. 1e+03;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
 |||||
 Db 9 RPQPPPP 15

RESULT 43
 ABM78986
 ID ABM78986 standard; protein; 1734 AA.
 XX
 AC ABM78986;
 XX
 DT 15-JAN-2004 (first entry)
 XX
 DE Modifier of beta-catenin (MBCAT) polypeptide GI 14149671.
 XX
 KW Modifier of beta-catenin; MBCAT; protein kinase; enzyme; cytostatic.
 XX
 OS Homo sapiens.

XX
 FH Key Location/Qualifiers
 FT Domain 512..785
 FT /note= "kinase domain"

XX WO2003074725-A2.
 XX
 PD 12-SEP-2003.
 XX
 PF 28-FEB-2003; 2003WO-US006294.
 XX
 PR 01-MAR-2002; 2002US-0361242P.
 XX
 PA (EXEL-) EXELIXIS INC.

XX
 PI Costa MA, Gendreau SB, Dora EG, Nicoll M, Heuer TS;
 XX
 DR WPI; 2003-756771/71.

XX
 PT Identifying a candidate beta-catenin pathway-modulating agent, for
 PT diagnosing and treating cancer, comprises contacting an assay system
 PT comprising a modifier of beta-catenin (MBCAT) polypeptide or nucleic acid
 PT with a test agent.

XX
 PS Claim 13; Page 79-86; 114pp; English.

XX
 CC The present sequence is that of human modifier of beta-catenin (MBCAT)
 CC polypeptide GI 14149671. MBCATs are kinase proteins with protein kinase
 CC domains. Genetic screens were designed to identify modifiers of the beta-
 CC catenin pathway in *Caenorhabditis elegans*, and the C10C6.1 gene was
 CC identified. Human orthologues, including the present sequence, were
 CC subsequently obtained. MBCAT genes and polypeptides can be used to
 CC identify MBCAT-modulating agents that are candidate therapeutic agents
 CC for treatment of disorders associated with defective or impaired beta-
 CC catenin and/or MBCAT function, such as an angiogenic, apoptotic or cell
 CC proliferation disorder, e.g. breast, colon, head and neck, kidney, lung,
 CC ovarian, prostate, skin or uterine cancer. MBCAT modulating agents
 CC include antisense oligomers and RNAi that repress gene expression or
 CC product activity. Transgenic non-human animals are useful as models of
 CC disease and disorders implicating defective beta-catenin function

XX
 SQ Sequence 1734 AA;

Query Match 100.0%; Score 45; DB 7; Length 1734;
 Best Local Similarity 100.0%; Pred. No. 1e+03;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
 |||||

Db 9 RPQPPPP 15

RESULT 44
ADG20497
ID ADG20497 standard; protein; 1734 AA.
XX ADG20497;
AC ADG20497;
XX 11-MAR-2004 (first entry)
XX Human MAST205 protein sequence SEQ ID NO:30.
DE diabetes; pre-diabetic; antidiabetic; insulin resistance; human; MAST205.
KW diabetes; pre-diabetic; antidiabetic; insulin resistance; human; MAST205.
XX Homo sapiens.
XX WO2003103601-A2.
XX 18-DEC-2003.
XX 05-JUN-2003; 2003WO-US018046.
XX 05-JUN-2002; 2002US-0386521P.
XX 05-JUN-2002; 2002US-0386527P.
XX 05-JUN-2002; 2002US-0386551P.
XX 06-JUN-2002; 2002US-0386423P.
XX 06-JUN-2002; 2002US-0386936P.
XX 06-JUN-2002; 2002US-0386954P.
XX 07-JUN-2002; 2002US-0387301P.
XX (META-) METABOLEX INC.
XX Allan B, Gregoire F, Lavan B, Moodie S, Waters S, Wong C;
PI WPI; 2004-062209/06.
XX N-PSDB; ADG20496.
XX Identifying an agent for treating a diabetic or pre-diabetic individual, comprises selecting an agent that modulates the expression or activity of a diabetes-related polypeptide encoded by a diabetes-related polynucleotide.
XX Claim 1; SEQ ID NO 30; 96pp; English.
XX The present invention describes a method for identifying an agent for treating a diabetic or pre-diabetic individual. The method comprises contacting an agent to a mixture comprising a polypeptide encoded by a nucleic acid that hybridizes under stringent conditions to a nucleic acid encoding any one of 10 sequences of 188-1734 amino acids (1) (see ADG20469, ADG20473, ADG20475, ADG20477, ADG20483, ADG20487, ADG20489, ADG20495, ADG20497 and ADG20501), and selecting an agent that modulates the expression or activity of a polypeptide or that binds to the polypeptide. Also described: (1) a method of treating a diabetic or pre-diabetic animal comprising administering to the animal a therapeutic amount of the agent identified by the method described above; (2) a method of introducing an expression cassette into a cell comprising introducing into the cell an expression cassette comprising a promoter operably linked to a polynucleotide encoding a polypeptide, where the polynucleotide hybridizes under stringent conditions to a nucleic acid encoding (1); (3) a method of diagnosing an individual who has type 2 diabetes or is pre-diabetic comprising detecting in a sample from the individual the level of a polypeptide or a polynucleotide encoding the polypeptide, where the polynucleotide hybridizes under stringent conditions to a nucleic acid encoding (1), and where a modulated level of the polypeptide or polynucleotide in the sample compared to a level of the polypeptide or polynucleotide in a lean individual or a previous sample from the individual indicates that the individual is diabetic or pre-diabetic; (4) an isolated nucleic acid encoding a polypeptide comprising a sequence of ADG20476 or ADG20494; (5) an expression cassette comprising the isolated nucleic acid; and (6) a host cell comprising the expression cassette. (1) has antidiabetic activity. The methods, polypeptides and polynucleotides are useful for identifying modulators of

CC the expression or activity of the polypeptides and polynucleotides. The modulators can be used for diagnosing and treating type 2 diabetes, pre-diabetes, or insulin resistance. The present sequence is used in the exemplification of the present invention.

CC the expression or activity of the polypeptides and polynucleotides. The modulators can be used for diagnosing and treating type 2 diabetes, pre-diabetes, or insulin resistance. The present sequence is used in the exemplification of the present invention.

XX Sequence 1734 AA;
Query Match 100.0%; Score 45; DB 8; Length 1734;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 RPQPPPP 7
Db 9 RPQPPPP 15

RESULT 45
ADQ15178
ID ADQ15178 standard; protein; 1734 AA.
XX ADQ15178;
AC ADQ15178;
XX 07-OCT-2004 (first entry)
XX Human cancer related polypeptide, 21708.
XX Cancer; angiogenic disorder; cytostatic; antiangiogenic; gene therapy; human.
XX Homo sapiens.
XX WO2004058153-A2.
XX 15-JUL-2004.
XX 16-DEC-2003; 2003WO-US040226.
XX 20-DEC-2002; 2002US-0435108P.
XX 23-DEC-2002; 2002US-0436443P.
XX 07-JAN-2003; 2003US-0438498P.
XX 31-JAN-2003; 2003US-0444370P.
XX 06-FEB-2003; 2003US-0446031P.
XX 11-MAR-2003; 2003US-0453635P.
XX 25-MAR-2003; 2003US-0457199P.
XX 10-APR-2003; 2003US-0462458P.
XX 30-APR-2003; 2003US-0466732P.
XX 08-MAY-2003; 2003US-0469184P.
XX 19-MAY-2003; 2003US-0471663P.
XX 03-JUN-2003; 2003US-0475472P.
XX 12-JUN-2003; 2003US-0478150P.
XX 23-JUN-2003; 2003US-0480631P.
XX 15-JUL-2003; 2003US-0487369P.
XX 29-JUL-2003; 2003US-0490866P.
XX 02-SEP-2003; 2003US-0499614P.
XX 09-OCT-2003; 2003US-0510081P.
XX 06-NOV-2003; 2003US-0517742P.
XX (MILL-) MILLENNIUM PHARM INC.
XX Lightcap ES, Eceedy JA, Hunter JJ, Macbeth KJ, Nestor MT;
XX WPI; 2004-525772/50.
XX N-PSDB; ADQ15177.
XX Identifying a compound capable of treating cancer or an angiogenic disorder by combining a compound to be tested with e.g., 15986, 2179 or 13249 polypeptide and detecting binding of the test compound to the polypeptide.
XX Claim 1; SEQ ID NO 136; 576pp; English.
XX The invention relates to a novel method for identifying a compound capable of treating cancer or an angiogenic disorder. The method

CC comprises: combining a compound to be tested with a cancer related
 CC polypeptide; and detecting binding of the test compound to the
 CC polypeptide. The invention further comprises: identifying a subject
 CC having or at risk for developing cancer; and treating a subject having a
 CC cancer or a cancer characterised by aberrant 15986, 2188, 20743, 9148,
 CC 9151, 9791, 44252, 14184, 42461, 8204, 7970, 25552, 21657, 26492, 2411,
 CC 15088, 1905, 28899, 63380, 33935, 10480, 12686, 25501, 17694, 15701,
 CC 53052, 49908, 21612, 38949, 6216, 46863, 9235, 2201, 6985, 9883, 12238,
 CC 18057, 21617, 39228, 49928, 54476, 62113, 64316, 12264, 32362, 58198,
 CC 2887, 3205, 8557, 9600, 9693, 44867, 53058, 55556, 57658, 2208, 10252,
 CC 10302, 14218, 33877, 10317, 10485, 25964, 14815, 1363, 1397, 14827,
 CC 21708, 3801, 64698, 2179 or 13249 cancer related polypeptide activity or
 CC nucleic acid expression. The method and compounds have cytostatic and
 CC antiangiogenic activities. The cancer related genes may be used in gene
 CC therapy to treat disorders. The cancer related genes may be used in gene
 CC compound capable of treating cancer or an angiogenic disorder. This
 CC sequence represents a human cancer related polypeptide used in the method
 CC of the invention.

XX SQ Sequence 1734 AA;

Query Match 100.0%; Score 45; DB 8; Length 1734;
 Best Local Similarity 100.0%; Pred. No. 1e+03;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
 DB 9 RPQPPPP 15

RESULT 46
 AAM49177
 ID AAM49177 standard; protein; 1798 AA.

XX AC AAM49177;

XX DT 14-JUN-2002 (first entry)

XX DE Human MAST205 (hMAST205).

XX KW Human; MAST205; hMAST205; PDZ domain; protocadherin LKC;
 KW microtubule-associated spermatid manchette serine/threonine kinase 205;
 KW cellular proliferation; cancer; tumour; drug screening; cytostatic;
 KW anticancer; enzyme.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers
 XX FT Domain 1032..1294

XX FT /label= PDZ_domain

XX FT /note= "Involved in binding the C-terminal 4 residues of
 XX FT protocadherin LKC (AAM49176, AAM49179)"

XX PN WO200198483-A1.

XX PD 27-DEC-2001.

XX PF 21-JUN-2001; 2001WO-JP005313.

XX PR 22-JUN-2000; 2000JP-00192693.

XX PR 09-FEB-2001; 2001JP-00034058.

XX PA (HELI-) HELIX RES INST.

XX PI Koga H, Takahashi N, Okazaki N, Masuho Y;

XX DR WPI; 2002-130792/17.

XX DR N-PSDB; ABL55198.

XX PT Anticancer function-carrying protocadherin-like protein, encoded gene and
 XX PT their variants, applicable in cancer diagnosis e.g. as target for
 XX PT evaluation of cancer cell proliferation and metastasis, and developing
 XX PT drugs for cancer.

XX PS Claim 4; Page 81-94; 101pp; Japanese.

XX CC The invention relates to a human protocadherin-like protein designated
 CC protocadherin LKC (AAM49176), to a human homologue of MAST205
 CC (microtubule-associated spermatid manchette serine/threonine kinase 205)
 CC designated hMAST205 (AAM49177) and to nucleic acids encoding them
 CC (ABL55191, ABL55198). Protocadherin LKC and hMAST205 are able to bind to
 CC each other via the interaction of the C-terminus of protocadherin LKC and
 CC the PDZ domain of hMAST205, and are involved in the regulation of
 CC cellular proliferation. The protein and nucleic acid sequences of the
 CC invention may be used in screening for drugs with anticancer activity.
 CC Additionally, protocadherin LKC may be used in the diagnosis of cancer as
 CC a target for the evaluation of cancer cell proliferation and metastasis,
 CC where the risk of cancer is assessed when there is a reduction in
 CC expression or a mutation in the protocadherin LKC gene. The present
 CC sequence represents hMAST205

XX SQ Sequence 1798 AA;

Query Match 100.0%; Score 45; DB 5; Length 1798;
 Best Local Similarity 100.0%; Pred. No. 1.1e+03;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
 DB 9 RPQPPPP 15

RESULT 47
 AAE16277
 ID AAE16277 standard; protein; 1798 AA.

XX AC AAE16277;

XX DT 26-MAR-2002 (first entry)

XX DE Human kinase PKIN-23 protein.

XX KW Human; kinase; PKIN-23; cancer; leukaemia; adenocarcinoma; osteoporosis;
 KW immune disorder; atherosclerosis; Crohn's disease; Hodgkin's disease;
 KW Acquired Immune Deficiency Syndrome; AIDS; Addison's disease; anaemia;
 KW allergy; asthma; adult respiratory distress syndrome; multiple sclerosis;
 KW autoimmune thyroiditis; bronchitis; diabetes mellitus; osteoarthritis;
 KW Good pasture's syndrome; Graves' disease; pancreatitis; psoriasis;
 KW rheumatoid arthritis; ulcerative colitis; cirrhosis; Cushing's syndrome;
 KW hepatitis; hypothyroidism; cerebral palsy; cataract; angina pectoris;
 KW cardiovascular disease; hypertension; vasculitis; myocarditis; obesity;
 KW congestive heart failure; ischaemic heart disease; lung tumour; gout;
 KW fatty liver; Niemann-Pick's disease; gene therapy; chromosome 13.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

XX FT Domain 512..785

XX FT /note= "Eukaryotic protein kinase domain"

XX FT Domain 1104..1191

XX FT /label= PDZ_domain

XX PN WO200196547-A2.

XX PD 20-DEC-2001.

XX PF 14-JUN-2001; 2001WO-US019444.

XX PR 15-JUN-2000; 2000US-0212073P.

XX PR 23-JUN-2000; 2000US-0213467P.

XX PR 30-JUN-2000; 2000US-0215651P.

XX PR 07-JUL-2000; 2000US-0216605P.

XX PR 13-JUL-2000; 2000US-0218372P.

XX PR 25-AUG-2000; 2000US-0228056P.

XX PA (INCY-) INCYTE GENOMICS INC.

XX Yue H, Lal P, Bandman O, Borowsky ML, Au-Young J, Lu Y;
 PI Gandhi AR, Tribouley CM, Wallia NK, Yao MG, Lu DAM, Greenwald SR;
 PI Ramkumar J, Griffin JA, Kearney L, Burford N, Nguyen DB, Tang YT;
 PI Baughn MR, He A, Thornton M, Hafalia A, Patterson C, Gururajan R;
 PI Lo TP, Khan F, Recipon SA, Azinzi Y, Policky JL, Ding L;
 PI Grether M, Elliott VS, Thangavelu K, Batra S, Ison CH;
 XX WPI; 2002-090207/12.
 DR N-PSDB; AAD26470.
 XX
 XX New polypeptides, useful for diagnosing, treating or preventing disorders
 PT of growth and development, cardiovascular and lipid, and diseases such as
 PT cancer, comprise human kinase polypeptides.
 XX
 PS Claim 1; Page 169-172; 197pp; English.
 XX
 CC The invention relates to human kinase PKIN proteins and their
 CC corresponding cDNAs. A composition containing PKIN agonist is useful for
 CC treating a disease or condition associated with decreased expression of
 CC PKIN and a composition comprising PKIN antagonist is useful for treating
 CC a disease or condition associated with overexpression of PKIN. The
 CC disorders include cancer (leukemia, adenocarcinoma, lymphoma, melanoma,
 CC myeloma, sarcoma, teratocarcinoma, Hodgkin's disease), immune disorder
 CC (Acquired Immune Deficiency Syndrome (AIDS), asthma, Addison's disease,
 CC atherosclerosis, anemia, allergies, adult respiratory distress syndrome,
 CC autoimmune thyroiditis, gout, bronchitis, Crohn's disease, diabetes
 CC mellitus, multiple sclerosis, Good pasture's syndrome, Graves' disease,
 CC osteoarthritis, osteoporosis, pancreatitis, psoriasis, Reiter's syndrome,
 CC rheumatoid arthritis, Sjogren's syndrome, uveitis, ulcerative colitis,
 CC bacterial, parasitic, fungal, viral, protozoal and helminthic infections)
 CC growth and development disorders (arteriosclerosis, cirrhosis, hepatitis,
 CC Cushing's syndrome, hypothyroidism, cerebral palsy, cataracts); cardio
 CC vascular disease (arteriovenous fistula, hypertension, vasculitis,
 CC aneurysms, congestive heart failure, angina pectoris, myocarditis,
 CC ischaemic heart disease, chronic bronchitis, lung tumours); lipid
 CC disorder (fatty liver, Fabry's disease, Niemann-Pick's disease,
 CC hypocholesterolemia, obesity). PKIN DNA is useful for assessing toxicity
 CC of a test compound and in gene therapy. The present sequence is human
 CC PKIN-23 protein. Human PKIN-23 gene is located on chromosome 13
 XX
 SQ Sequence 1798 AA;
 Query Match 100.0%; Score 45; DB 5; Length 1798;
 Best Local Similarity 100.0%; Pred. No. 1.1e+03;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 RPQPPPP 7
 Db 9 RPQPPPP 15
 RESULT 48
 ADJ96613
 ID ADJ96613 standard; protein; 1798 AA.
 AC ADJ96613;
 XX
 XX 06-MAY-2004 (first entry)
 DT
 XX Human microtubule associated ser/thr protein kinase MAST205 SeqID 70.
 DE
 XX kinase; human; SNP; single nucleotide polymorphism;
 KW tyrosine protein kinase; serine/threonine protein kinase; PTK; STK;
 KW gene therapy; cancer; immune-related disease; cardiovascular disease;
 KW brain; neuronal associated disease; metabolic; inflammatory disorder;
 KW cystostatic; neuroprotective; immunomodulator; antiinflammatory; enzyme;
 KW microtubule associated ser/thr protein kinase; MAST205.
 XX
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FH Misc-difference 9
 FT

FT
 FT /note= "Wild type Arg substituted for STOP by single
 nucleotide polymorphism"
 FT Misc-difference 1551
 FT /note= "Wild type Asp substituted for Gly by single
 nucleotide polymorphism"
 FT Misc-difference 1768
 FT /note= "Wild type Ser substituted for Phe by single
 nucleotide polymorphism"
 FT
 XX WO2004006838-A2.
 PN
 XX 22-JAN-2004.
 PD
 XX 15-JUL-2003; 2003WO-US021730.
 PF
 XX 15-JUL-2002; 2002US-0395632P.
 PR
 XX (SUGEN-) SUGEN INC.
 PA
 XX Whyte D, Manning G, Caenepeel S;
 PI WPI; 2004-122753/12.
 DR N-PSDB; ADJ96547.
 XX
 PT New nucleic acid molecule encoding a kinase polypeptide, useful for
 PT preparing a composition for treating diseases or disorders, e.g., cancer,
 PT or neurological, immunological or inflammatory disorders.
 XX
 PS Claim 1; SEQ ID NO 70; 366pp; English.
 XX
 CC This invention relates to a novel isolated, enriched or purified nucleic
 CC acid molecule that encodes a kinase polypeptide. Specifically, it relates
 CC to human tyrosine and serine/threonine protein kinases (PTK's and STK's),
 CC as well as protein kinase-like enzymes. The present invention describes
 CC screening methods to identify agonists, antagonists and antibodies that
 CC can be used to modulate the activity or function of the mammalian kinase
 CC enzymes. As such, these compositions can be used for gene therapy
 CC purposes to treat diseases or disorders including cancer, immune-related
 CC diseases, cardiovascular disease, brain or neuronal associated disease,
 CC metabolic and inflammatory disorders. Accordingly, they exhibit
 CC cytosstatic, neuroprotective, immunomodulator and antiinflammatory
 CC activities. This polypeptide sequence is a human kinase protein sequence
 CC of the invention.
 XX
 SQ Sequence 1798 AA;
 Query Match 100.0%; Score 45; DB 8; Length 1798;
 Best Local Similarity 100.0%; Pred. No. 1.1e+03;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 RPQPPPP 7
 Db 9 RPQPPPP 15
 RESULT 49
 ADX91705
 ID ADX91705 standard; protein; 1805 AA.
 XX
 XX ADX91705;
 AC
 XX 21-APR-2005 (first entry)
 DT
 XX Plant full length insert polypeptide seqid 54369.
 DE
 XX plant protectant; plant growth regulant; gene therapy; plant;
 KW recombinant DNA construct; physical array; plant breeding marker;
 KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
 KW extreme osmotic condition; pathogen tolerance; pest tolerance;
 KW growth rate; cell cycle pathway; disease resistance;
 KW galactomannan production; lignin production; plant growth regulator;
 KW yield; plant growth; plant development; seed oil; protein yield;
 KW protein content.

XX OS Unidentified.
 XX PN US2004034888-A1.
 XX PD 19-FEB-2004.
 XX PF 28-APR-2003; 2003US-00425114.
 XX PR 06-MAY-1999; 99US-00304517.
 XX PR 05-NOV-2001; 2001US-00985678.
 XX PA (LIU//) LIU J.
 XX PA (ZHOU//) ZHOU Y.
 XX PA (KOVA//) KOVALIC D K.
 XX PA (SCRE//) SCREEN S E.
 XX PA (TABA//) TABASKA J E.
 XX PA (CAOY//) CAO Y.
 XX PI Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
 XX WPI; 2004-180133/17.
 XX DR
 XX XX New recombinant DNA construct, useful for improving plant tolerance to
 XX PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
 XX PT pests, for conferring increased resistance to plant disease, or for
 XX PT improving yield.
 XX PS Claim 1; SEQ ID NO 54369; 15pp; English.
 XX CC The invention describes a recombinant DNA construct comprising a
 XX CC polynucleotide consisting of a sequence encoding an amino acid sequence
 XX CC available in electronic form from the US patent office at
 XX CC ftp.segdata.uspto.gov/sequence.html?DocID:2004034888. The polynucleotide
 XX CC of the invention are also useful in physical arrays of molecules and as
 XX CC plant breeding markers. The recombinant DNA construct is useful for
 XX CC improving plant tolerance to cold, heat, drought, herbicides, extreme
 XX CC osmotic conditions, pathogens or pests, for manipulating growth rate in
 XX CC plant cells by modification of the cell cycle pathway, for conferring
 XX CC increased resistance to plant disease, for producing galactomannan,
 XX CC lignin or plant growth regulators, for increasing the rate of homologous
 XX CC recombination in plants, for improving yield by modification of
 XX CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
 XX CC or by providing improved plant growth and development under at least one
 XX CC stress condition or for modifying seed oil or protein yield and/or
 XX CC content. This is the amino acid sequence of a plant full length insert
 XX CC polypeptide that can be used in the recombinant DNA construct of the
 XX CC invention.
 XX PS Sequence 1805 AA;
 XX CC
 XX CC Query Match 100.0%; Score 45; DB 8; Length 1805;
 XX CC Best Local Similarity 100.0%; Pred. No. 1.1e+03;
 XX CC Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX CC
 XX CC 1 RPQPPPP 7
 XX CC 17 RPQPPPP 23
 XX CC
 XX CC RESULT 50
 XX CC ADH47759
 XX CC ID ADH47759 standard; protein; 4952 AA.
 XX CC AC ADH47759;
 XX CC XX
 XX CC 25-MAR-2004 (first entry)
 XX CC XX
 XX CC NOV18 protein, SEQ ID 56.
 XX CC DE
 XX CC Antidiabetic; anorectic; cardiant; hypotensive; antiarteriosclerotic;
 XX CC KW anorectic; virucide; antibacterial; fungicide; protozoacide; nootropic;
 XX CC KW neuroprotective; antiparkinsonian; anticonvulsant; osteopathic;

KW antiarthritic; antiinflammatory; dermatological; antiasthmatic;
 KW antilipaemic; Gene therapy; human; metabolic disorder; diabetes; obesity;
 KW viral infection; bacterial infection; fungal infection;
 KW helminthic infection; protozoal infection; anorexia; cancer;
 KW cardiovascular disease; neurodegenerative disorder; Alzheimer's disease;
 KW Parkinson's disease; epilepsy; immune disorder; haematopoietic disorder;
 KW inflammatory skin disorder; asthma; dyslipidaemia; NOV18;
 KW ARL-like protein; chromosome 12q12-q14.
 XX OS Homo sapiens.
 XX PN WO200268647-A2.
 XX PD 06-SEP-2002.
 XX PF 16-JAN-2002; 2002WO-US001311.
 XX PR 16-JAN-2001; 2001US-0261376P.
 XX PR 18-JAN-2001; 2001US-0262454P.
 XX PR 18-JAN-2001; 2001US-0262587P.
 XX PR 31-JAN-2001; 2001US-0265530P.
 XX PR 14-FEB-2001; 2001US-0268595P.
 XX PR 28-FEB-2001; 2001US-0272409P.
 XX PR 16-MAR-2001; 2001US-0276777P.
 XX PR 17-MAY-2001; 2001US-0291672P.
 XX PR 27-SEP-2001; 2001US-0325306P.
 XX PR 18-OCT-2001; 2001US-033036P.
 XX PR 09-NOV-2001; 2001US-0345202P.
 XX PA (CURA-) CURAGEN CORP.
 XX PI Padigar M, Alsobrook JP, Colman SD, Spytek KA, Boldog F;
 XX PI Vernet CAM, Li L, Shenoy S, Casman S, Guo X, Edinger S;
 XX PI MacDougall J, Malyankar U, Patturajan M, Shimkets RA, Pena C;
 XX PI Tchernev V, Zerhusen BD, Millett I, Miller C, Lepley DM, Smithson G;
 XX PI Baumgartner J, Herrmann J, Peyman JA, Gorman L, Mezes P, Kekuda R;
 XX PI Taupier RJ, Gerlach V, Grosse WM, Liu X, Ellerman K, Rothenberg M;
 XX PI Stone DJ, Burgess CE;
 XX WPI; 2002-698671/75.
 XX N-PSDB; ADH47758.
 XX CC New isolated NOVX polypeptides and polynucleotides, useful for
 XX CC preventing, diagnosing or treating NOVX-associated disorders e.g.
 XX CC osteoarthritis, obesity, atherosclerosis, cancer, Parkinson's disease,
 XX CC asthma, or infections.
 XX PS Claim 1; Page 159-160; 380pp; English.
 XX CC The present invention relates to novel proteins (I) referred to as NOVX,
 XX CC where X is any number from 1 to 18, and their coding sequences (II). The
 XX CC proteins and their coding sequences are useful in the manufacture of a
 XX CC medicament for treating a syndrome associated with a human disease,
 XX CC preferably a NOVX-associated disorder such as metabolic disorders,
 XX CC diabetes, obesity, infectious diseases (viral, bacterial, fungal,
 XX CC helminthic, and protozoal), anorexia, cancer, cardiovascular diseases
 XX CC (hypertension, atherosclerosis), neurodegenerative disorders, Alzheimer's
 XX CC disease, Parkinson's disease, epilepsy, immune disorders
 XX CC (osteoarthritis), haematopoietic disorders, inflammatory skin disorders,
 XX CC asthma, and various dyslipidaemias. NOV18 comprises a ARL-like protein
 XX CC and maps to chromosome 12q12-q14.
 XX SQ Sequence 4952 AA;
 XX CC
 XX CC Query Match 100.0%; Score 45; DB 5; Length 4952;
 XX CC Best Local Similarity 100.0%; Pred. No. 2.5e+03;
 XX CC Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX CC
 XX CC 1 RPQPPPP 7
 XX CC 1806 RPQPPPP 1812
 XX CC

Search completed: April 6, 2006, 09:32:23
Job time : 109.289 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: April 6, 2006, 09:35:13 ; Search time 12.7105 Seconds
(without alignments)
52.989 Million cell updates/sec

Title: US-10-632-388-305

Perfect score: 45

Sequence: 1 RQPPPPP 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Database :

1: PIR 80:*

2: PIR1:*

3: PIR3:*

4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	45	100.0	125	2 A53692	synapsin I - mouse
2	45	100.0	691	2 A25704	synapsin I - rat
3	45	100.0	704	2 A30411	synapsin Ia - rat
4	45	100.0	705	2 A35363	synapsin I splice
5	45	100.0	706	2 E30411	synapsin Ia - bovi
6	45	100.0	4957	2 T03455	ALR protein - huma
7	45	100.0	5262	2 T03455	ALR protein - huma
8	42	93.3	139	2 T05947	hypothetical prote
9	42	93.3	320	2 AE2842	conserved hypotet
10	42	93.3	359	2 F97619	hypothetical prote
11	42	93.3	645	2 A71416	hypothetical prote
12	42	93.3	736	2 I51691	dishevelled homolo
13	42	93.3	1278	2 T03188	Niemann-Pick C dis
14	42	93.3	1745	2 A46431	tight junction-ass
15	41	91.1	159	2 A83295	hypothetical prote
16	41	91.1	248	1 JQ1882	infected cell prot
17	41	91.1	320	1 A39724	homeotic protein H
18	41	91.1	334	1 A41786	mRNA-binding prote
19	41	91.1	634	2 T00388	hypothetical prote
20	41	91.1	780	2 T00366	hypothetical prote
21	41	91.1	972	2 T47922	anthranilate phosph
22	41	91.1	979	2 A35913	regulatory factor
23	41	91.1	1039	2 T30351	mucin-like protein
24	41	91.1	1297	2 A57278	fibrillin-2 precur
25	41	91.1	2918	2 A54105	fibrillin-2 precur
26	41	91.1	3144	2 A46068	Huntington disease
27	40	88.9	70	2 A96850	protein F2401.6 [i
28	40	88.9	196	2 B48232	cysteine-rich exte
29	40	88.9	209	2 A48232	cysteine-rich exte

30	40	88.9	215	2 S34163	homeotic protein H
31	40	88.9	221	2 T07176	extensin homolog -
32	40	88.9	242	2 S54156	extensin-like prot
33	40	88.9	302	2 C86480	33.2K hypothetical
34	40	88.9	303	2 JQ1386	hypothetical 33K p
35	40	88.9	312	2 B29350	env polyprotein, r
36	40	88.9	334	2 E70519	probable oxidoredu
37	40	88.9	334	2 T33836	hypothetical prote
38	40	88.9	346	2 JCS833	transcription fact
39	40	88.9	349	2 T49791	hypothetical prote
40	40	88.9	356	1 VCVWB6	env polyprotein pr
41	40	88.9	356	2 S31428	envelope protein -
42	40	88.9	408	1 VCMVSR	env polyprotein -
43	40	88.9	408	2 S50899	betab inhibin prec
44	40	88.9	409	2 A21170	gp55 protein - Fri
45	40	88.9	411	2 S34164	homeotic protein H
46	40	88.9	413	2 T04520	hypothetical prote
47	40	88.9	416	2 S27198	homeotic protein H
48	40	88.9	416	2 T34279	hypothetical prote
49	40	88.9	417	2 S47539	homeotic protein H
50	40	88.9	422	2 A83184	probable protein m
51	40	88.9	456	2 S23104	choline kinase - h
52	40	88.9	469	2 I37451	HBF-G2 (HFK-2) pro
53	40	88.9	476	2 A54743	transcription fact
54	40	88.9	480	2 JH0672	brain factor 1 pro
55	40	88.9	481	2 T27975	hypothetical prote
56	40	88.9	489	2 T26069	hypothetical prote
57	40	88.9	495	1 S31223	transcription fact
58	40	88.9	533	1 S56652	calcium-dependent
59	40	88.9	546	2 T02029	DNA-binding protei
60	40	88.9	549	2 C87719	protein R119.6 [im
61	40	88.9	588	2 T45564	hypothetical prote
62	40	88.9	610	2 S35049	mucin JER57 - huma
63	40	88.9	610	2 S55643	glycoprotein L - e
64	40	88.9	617	2 I56530	gene VGF protein -
65	40	88.9	636	1 VCVWFS	env polyprotein -
66	40	88.9	640	1 VCVWM1	env polyprotein -
67	40	88.9	640	1 VCMVRV	env polyprotein pr
68	40	88.9	655	2 AD2422	DNA polymerase III
69	40	88.9	655	2 S40521	FKHR protein - hum
70	40	88.9	656	2 B49423	semaphorin I - fru
71	40	88.9	669	2 I38029	matrix metalloprot
72	40	88.9	756	2 T27642	hypothetical prote
73	40	88.9	785	2 S75819	general secretion
74	40	88.9	839	2 T04859	extensin homolog F
75	40	88.9	963	2 A55926	DNA binding protei
76	40	88.9	1026	1 A40315	maternal effect pr
77	40	88.9	1174	2 T43051	protein kinase C (
78	40	88.9	1201	2 A57369	anillin - fruit fl
79	40	88.9	1213	2 S16356	ovo protein - fru
80	40	88.9	1234	2 T30160	hypothetical prote
81	40	88.9	1240	2 JCS209	insulin receptor s
82	40	88.9	1242	2 JS0670	insulin receptor s
83	40	88.9	1316	2 T00381	KIA0633 protein -
84	40	88.9	1403	2 S24548	homeotic protein p
85	40	88.9	1460	1 EDBEIF	immediate-early pr
86	40	88.9	1541	2 T02831	AAA protein L4171.
87	40	88.9	1840	2 T30250	GTP protein - mous
88	40	88.9	2649	2 T51023	hypothetical prote
89	39	86.7	126	2 H86347	hypothetical prote
90	39	86.7	147	2 S37485	gene msl protein
91	39	86.7	148	2 S39206	proline-rich prote
92	39	86.7	148	2 E75300	conserved hypotet
93	39	86.7	161	2 T04981	hypothetical prote
94	39	86.7	161	2 H85120	probable proline-r
95	39	86.7	198	2 T35752	hypothetical prote
96	39	86.7	199	2 S14981	extensin class I (
97	39	86.7	213	2 C86575	hypothetical prote
98	39	86.7	213	2 B72049	hypothetical prote
99	39	86.7	219	2 T10563	proline-rich prote
100	39	86.7	219	2 T27105	hypothetical prote
101	39	86.7	248	2 F91197	BapF protein [impo
102	39	86.7	248	2 B86044	espF [imported] -

103	39	86.7	278	2	A49266	fas ligand - rat	176	38	84.4	633	2	S62057	proline-rich prote
104	39	86.7	279	2	A53062	Fas ligand - mouse	177	38	84.4	982	1	GNLJH2	pol polyprotein -
105	39	86.7	281	1	T38707	Fas ligand - human	178	38	84.4	1091	2	S33596	protein-tyrosine k
106	39	86.7	303	2	T06435	ribonuclease S5 ho	179	38	84.4	1122	2	T47424	hypothetical prote
107	39	86.7	336	2	C48423	homeotic protein e	180	38	84.4	1338	2	T30565	MAP kinase kinase
108	39	86.7	348	2	T47494	hypothetical prote	181	38	84.4	1474	2	T18281	hypothetical prote
109	39	86.7	373	2	T51916	hypothetical prote	182	38	84.4	1776	2	G86280	protein TSE21.13 {
110	39	86.7	382	2	B86575	hypothetical prote	183	38	84.4	1872	2	T00339	hypothetical prote
111	39	86.7	382	2	A72049	hypothetical prote	184	38	84.4	2584	2	T24158	hypothetical prote
112	39	86.7	383	2	A82295	hypothetical prote	185	38	84.4	2606	2	T24157	hypothetical prote
113	39	86.7	393	2	T32533	hypothetical prote	186	37	82.2	75	2	S14973	extensin class I (
114	39	86.7	429	1	FOLJCN	gag polyprotein -	187	37	82.2	104	2	H96000	hypothetical prote
115	39	86.7	429	1	FOLJGH	probable gag polyp	188	37	82.2	107	2	E72710	hypothetical prote
116	39	86.7	429	2	S06073	gag polyprotein -	189	37	82.2	142	2	A11967	hypothetical prote
117	39	86.7	439	2	S51939	chitinase (EC 3.2.	190	37	82.2	143	2	D83148	hypothetical prote
118	39	86.7	458	2	S22697	extensin - Volvox	191	37	82.2	145	2	T22693	hypothetical prote
119	39	86.7	508	2	T45867	hypothetical prote	192	37	82.2	160	2	D48232	cysteine-rich exte
120	39	86.7	516	2	T00974	probable Sfr16 prot	193	37	82.2	161	2	E48232	cysteine-rich exte
121	39	86.7	574	2	T43556	Wiskott-Aldrich sy	194	37	82.2	161	2	F72593	hypothetical prote
122	39	86.7	574	2	T38819	wiskott-aldrich sy	195	37	82.2	165	2	C48232	cysteine-rich exte
123	39	86.7	610	2	T06690	galactonolactone d	196	37	82.2	168	2	T48343	hypothetical prote
124	39	86.7	643	2	T27429	hypothetical prote	197	37	82.2	168	2	T48375	transcription co-a
125	39	86.7	653	2	T08984	auxin response fac	198	37	82.2	215	2	B75281	peptidyl-tRNA hydr
126	39	86.7	715	1	TNBE77	77K alpha trans-in	199	37	82.2	235	2	T31424	C-terminal domain-
127	39	86.7	742	2	F84643	hypothetical prote	200	37	82.2	266	1	C1FGL	calpain (EC 3.4.22
128	39	86.7	760	2	T06291	extensin homolog T	201	37	82.2	266	1	CIHBL	calpain (EC 3.4.22
129	39	86.7	786	2	T49414	related to ahmpl p	202	37	82.2	268	1	CIHUL	calpain (EC 3.4.22
130	39	86.7	788	2	T25061	hypothetical prote	203	37	82.2	271	2	A70415	molybdopter in bios
131	39	86.7	814	2	G02390	disintegrin-like m	204	37	82.2	294	2	T34537	hypothetical prote
132	39	86.7	868	2	T20239	hypothetical prote	205	37	82.2	329	2	T01333	hypothetical prote
133	39	86.7	872	2	T50369	probable serine/th	206	37	82.2	354	2	A48931	transmembrane glyc
134	39	86.7	894	2	F84870	hypothetical prote	207	37	82.2	361	2	S19552	potassium channel
135	39	86.7	969	2	T15446	hypothetical prote	208	37	82.2	367	2	T01751	gibberellin 20-oxi
136	39	86.7	992	1	GNWR78	structural polypro	209	37	82.2	376	1	Q0BEW2	UL53 protein - hum
137	39	86.7	1071	2	S48378	probable membrane	210	37	82.2	387	2	T21370	hypothetical prote
138	39	86.7	1111	2	T23047	hypothetical prote	211	37	82.2	390	2	AC2169	molybdopter in bios
139	39	86.7	1147	2	T42627	ADP-ribosylation f	212	37	82.2	391	2	T52532	molybdopter in synt
140	39	86.7	1201	2	G86441	unknown protein [i	213	37	82.2	392	2	S76179	hypothetical prote
141	39	86.7	1201	2	T29329	hypothetical prote	214	37	82.2	392	2	G70594	probable MOEB- hom
142	39	86.7	1247	2	T45743	hypothetical prote	215	37	82.2	396	1	S17929	transcription init
143	39	86.7	1307	2	S53412	probable membrane	216	37	82.2	400	2	S58222	PQ-rich protein -
144	39	86.7	1443	2	S05979	steroid hormone re	217	37	82.2	425	2	S33045	hypothetical prote
145	39	86.7	1575	2	S68448	synaptotjanin, 170K	218	37	82.2	440	1	XXHUN	phosphatidylcholin
146	39	86.7	1594	2	T18276	protein-tyrosine k	219	37	82.2	440	2	JC1502	phosphatidylcholin
147	39	86.7	1599	2	T15854	hypothetical prote	220	37	82.2	473	2	S50755	hypothetical prote
148	39	86.7	1997	1	S12050	protein-tyrosine-p	221	37	82.2	481	2	E83062	probable outer mem
149	39	86.7	3051	2	S42373	hypothetical prote	222	37	82.2	487	2	E83120	hypothetical prote
150	39	86.7	7962	2	A30996	elastic titin - hu	223	37	82.2	533	2	D96554	tyrosine phosphor
151	38	84.4	106	2	A30996	orf-X protein - si	224	37	82.2	621	2	D96554	hypothetical prote
152	38	84.4	112	1	ASLJX3	vpu protein - simi	225	37	82.2	646	2	S55048	protein-tyrosine k
153	38	84.4	112	1	ASLJX3	vpu protein - huma	226	37	82.2	655	2	A81533	hypothetical prote
154	38	84.4	112	1	ASLJST	vpu protein - simi	227	37	82.2	679	2	A42073	potassium channel
155	38	84.4	112	1	ASLJX2	vpu protein - huma	228	37	82.2	681	2	F85062	hypothetical prote
156	38	84.4	112	2	S12155	vpx protein - huma	229	37	82.2	695	2	A87494	topoisomerase IV,
157	38	84.4	112	2	S53094	vpx protein - huma	230	37	82.2	751	2	T15403	hypothetical prote
158	38	84.4	112	2	S03066	gene X protein - h	231	37	82.2	760	2	F86387	hypothetical prote
159	38	84.4	152	2	T00877	hypothetical prote	232	37	82.2	769	2	I56546	probable Pco kinas
160	38	84.4	156	1	GNVQL2	genome-linked prot	233	37	82.2	793	1	KXRTF	Shaw type potassu
161	38	84.4	156	1	GNVQL2	genome-linked prot	234	37	82.2	798	2	T21369	furin (EC 3.4.21.7
162	38	84.4	156	1	GNVQL2	genome-linked prot	235	37	82.2	806	1	JN0612	hypothetical prote
163	38	84.4	219	2	GNVQWA	genome-linked prot	236	37	82.2	807	1	I51153	protein kinase B-r
164	38	84.4	280	2	T50583	ABC-type transport	237	37	82.2	815	2	T15402	hypothetical prote
165	38	84.4	304	1	A60472	uracil-DNA glycosy	238	37	82.2	897	2	A54696	EGF receptor subst
166	38	84.4	425	2	T16433	hypothetical prote	239	37	82.2	908	2	T25035	hypothetical prote
167	38	84.4	450	2	A53771	transcription fact	240	37	82.2	1006	2	G86292	hypothetical prote
168	38	84.4	457	2	JC6551	chitinase (EC 3.2.	241	37	82.2	1160	2	F88369	protein unc-52 [im
169	38	84.4	501	2	A42030	alpha-globin trans	242	37	82.2	1255	1	A24571	protein-tyrosine k
170	38	84.4	502	2	C56205	transcription fact	243	37	82.2	1313	2	F96673	hypothetical prote
171	38	84.4	502	2	B42030	alpha-globin trans	244	37	82.2	2115	2	S38480	nonstructural prot
172	38	84.4	547	2	A82965	hypothetical prote	245	37	82.2	2295	2	C88369	hypothetical prote
173	38	84.4	547	2	C82655	hypothetical prote	246	37	82.2	2783	1	A41948	alpha-fetoprotein
174	38	84.4	618	2	T05518	hypothetical prote	247	37	82.2	3078	2	T28432	variant-specific s
175	38	84.4	626	1	NDBE61	exonuclease (EC 3.	248	37	82.2	3375	2	T19821	hypothetical prote

249	37	82.2	26926	1	I38344	titin, cardiac mus	322	36	80.0	816	2	F31277	regulatory protein
250	36	80.0	75	2	T26916	hypothetical prote	323	36	80.0	828	2	T06133	hypothetical prote
251	36	80.0	80	2	P00477	pistil extensin-li	324	36	80.0	846	1	Q0BEC3	HQRf1 protein - hu
252	36	80.0	98	2	T42171	hypothetical prote	325	36	80.0	882	2	S41034	hypothetical prote
253	36	80.0	117	2	D40750	proline-rich prote	326	36	80.0	893	2	G88551	protein M01A8.2 [i
254	36	80.0	122	2	PQ0452	extensin-like prot	327	36	80.0	903	2	T09143	alpha-glucosidase
255	36	80.0	151	2	JQ1686	extensin-like prote	328	36	80.0	913	2	JCS463	alpha-glucosidase
256	36	80.0	160	2	C72501	hypothetical prote	329	36	80.0	981	1	POWVGM	gag-abl polyprotei
257	36	80.0	191	2	JQ1919	hypothetical prote	330	36	80.0	1000	2	S44898	ZK1236.3 protein -
258	36	80.0	211	2	S28304	hypothetical prote	331	36	80.0	1003	2	T34066	hypothetical prote
259	36	80.0	212	2	B36298	proline-rich prote	332	36	80.0	1008	2	T04462	hypothetical prote
260	36	80.0	229	1	W4WLB	E4 protein - human	333	36	80.0	1010	2	A33509	vinculin - Caenorh
261	36	80.0	231	2	B83031	conserved hypochet	334	36	80.0	1013	2	T33470	hypothetical prote
262	36	80.0	232	2	T10646	hypothetical prote	335	36	80.0	1066	2	T10108	vinculin - mouse
263	36	80.0	235	1	WZBE23	gene 23 protein -	336	36	80.0	1069	2	D85383	hypothetical prote
264	36	80.0	239	2	T16159	hypothetical prote	337	36	80.0	1072	2	T50949	verprolin related
265	36	80.0	240	2	F87537	TonB protein, prob	338	36	80.0	1100	2	T21544	hypothetical prote
266	36	80.0	245	1	W4WLB5	E4 protein - human	339	36	80.0	1123	2	A33962	kinase-related tra
267	36	80.0	245	1	W4WLB5	E4 protein - human	340	36	80.0	1130	1	TVHUA	protein-tyrosine k
268	36	80.0	249	2	T04939	hypothetical prote	341	36	80.0	1134	1	A35955	meta-vinculin - hu
269	36	80.0	278	2	B39066	proline-rich prote	342	36	80.0	1135	1	A29997	high molecular mas
270	36	80.0	281	2	S65692	homeoprotein, IDX-	343	36	80.0	1151	2	T18535	synGAP-b1 protein
271	36	80.0	283	2	S42634	hypothetical prote	344	36	80.0	1166	2	T13958	probable membrane
272	36	80.0	296	2	H87511	hypothetical prote	345	36	80.0	1219	2	S54570	structural polypro
273	36	80.0	299	2	G70784	probable mmpS3 pro	346	36	80.0	1245	1	VHWVB	myosin I myoA - Em
274	36	80.0	304	1	W4WLB4	E4 protein - human	347	36	80.0	1245	1	VHWVB2	structural polypro
275	36	80.0	309	2	S10889	proline-rich prote	348	36	80.0	1249	2	A56511	myosin I myoA - Em
276	36	80.0	309	2	B87576	hypothetical prote	349	36	80.0	1249	2	T14270	Ras-GTPase activat
277	36	80.0	312	2	A56911	TRADD protein - hu	350	36	80.0	1260	2	T01334	hypothetical prote
278	36	80.0	336	2	S29507	metavinculin - pig	351	36	80.0	1262	1	T13353	protein stn-B - fr
279	36	80.0	340	2	T19105	phosphate carrier	352	36	80.0	1293	2	T14259	ras GTPase-activat
280	36	80.0	342	2	T26677	hypothetical prote	353	36	80.0	1367	2	T13703	tamA protein - fru
281	36	80.0	351	2	B34768	ORF5 protein - Orf	354	36	80.0	1505	2	S28079	DNA-directed DNA p
282	36	80.0	356	2	A34350	phosphate transpor	355	36	80.0	1596	2	A35927	190K DNA-binding p
283	36	80.0	361	2	B53737	phosphate carrier	356	36	80.0	1733	1	B45344	probable nuclear a
284	36	80.0	361	2	D53737	phosphate carrier	357	36	80.0	1801	2	T26774	hypothetical prote
285	36	80.0	362	2	A53737	phosphate carrier	358	36	80.0	1958	1	B40505	hypothetical prote
286	36	80.0	362	2	A53737	phosphate carrier	359	36	80.0	2148	2	A50801	insulin receptor -
287	36	80.0	363	2	B84565	probable spliceoso	360	36	80.0	2182	2	T28634	variant-specific s
288	36	80.0	363	2	T16755	hypothetical prote	361	36	80.0	2715	2	T13049	eyellid - fruit fly
289	36	80.0	367	2	I39172	cyclin-dependent k	362	36	80.0	3005	2	T08841	polyprotein - dour
290	36	80.0	379	2	A47659	farnesyl-protein t	363	35	77.8	67	2	S14971	extensin class I (
291	36	80.0	379	2	T19069	hypothetical prote	364	35	77.8	76	2	E97763	protein transport
292	36	80.0	387	2	F96798	hypothetical prote	365	35	77.8	82	2	T30135	hypothetical prote
293	36	80.0	388	2	JCS437	spliceosome-associ	366	35	77.8	90	2	S14972	extensin class I (
294	36	80.0	392	1	PIHUB6	salivary proline-r	367	35	77.8	92	2	B48831	vitelline membrane
295	36	80.0	392	2	T15755	hypothetical prote	368	35	77.8	100	2	T17126	hypothetical prote
296	36	80.0	392	2	A86206	hypothetical prote	369	35	77.8	107	2	S20693	hypothetical prote
297	36	80.0	393	2	PQ0479	pistil extensin-li	370	35	77.8	107	2	C87312	hypothetical prote
298	36	80.0	397	2	S09813	hypothetical prote	371	35	77.8	108	2	T05231	hypothetical prote
299	36	80.0	403	2	S52796	prpL2 protein - hu	372	35	77.8	109	2	AG2992	conserved hypochet
300	36	80.0	409	2	T24543	hypothetical prote	373	35	77.8	109	2	C98291	hypothetical prote
301	36	80.0	426	2	JQ1696	pistil extensin-li	374	35	77.8	131	2	B96566	hypothetical prote
302	36	80.0	432	2	AB2222	twitching motility	375	35	77.8	132	2	S14970	extensin class I (
303	36	80.0	444	2	E83802	hypothetical prote	376	35	77.8	132	2	H96708	hypothetical prote
304	36	80.0	464	2	C84686	probable AP2 domai	377	35	77.8	134	2	JCS572	proline-rich prote
305	36	80.0	465	2	A44498	radial spoke prote	378	35	77.8	138	2	C96734	hypothetical prote
306	36	80.0	468	1	POWVMU	gag polyprotein -	379	35	77.8	139	2	T46186	hypothetical prote
307	36	80.0	538	1	POWVM	gag polyprotein -	380	35	77.8	142	2	B41132	collagen-related p
308	36	80.0	538	1	POWVM	gag polyprotein -	381	35	77.8	144	2	E86364	hypothetical prote
309	36	80.0	547	2	T02274	hypothetical prote	382	35	77.8	148	2	D72652	hypothetical prote
310	36	80.0	549	2	I37386	fas-activated seri	383	35	77.8	148	2	AD2681	transcription regu
311	36	80.0	557	2	S23429	protein kinase ERK	384	35	77.8	148	2	B97463	probable repressor
312	36	80.0	560	2	I59302	brain specific Na+	385	35	77.8	149	2	A41132	collagen-related p
313	36	80.0	564	2	T45866	hypothetical prote	386	35	77.8	154	2	PQ0476	pistil extensin-li
314	36	80.0	602	2	T45760	hypothetical prote	387	35	77.8	156	2	C84556	probable b2IP tran
315	36	80.0	606	1	UZAD12	terminal protein p	388	35	77.8	159	2	A72680	hypothetical prote
316	36	80.0	618	2	JC4366	transcription fact	389	35	77.8	163	2	A29356	hydroxyproline-ric
317	36	80.0	651	2	T21175	hypothetical prote	390	35	77.8	164	2	G71427	hypothetical prote
318	36	80.0	671	2	T02504	hypothetical prote	391	35	77.8	166	2	T09593	cIC protein, cold-
319	36	80.0	721	2	C84677	probable membrane	392	35	77.8	168	2	T20708	hypothetical prote
320	36	80.0	778	2	T30430	hypothetical prote	393	35	77.8	169	2	T07623	extensin homolog H
321	36	80.0	782	2	T48722	hypothetical prote	394	35	77.8	172	2	A39458	carotene biosynthe

395	35	77.8	176	2	B85355	hypothetical prote	468	35	77.8	385	2	S78100	MAPK-activated pro
396	35	77.8	181	2	S14974	extensin class I (469	35	77.8	387	2	C96720	hypothetical prote
397	35	77.8	182	2	T30760	hypothetical prote	470	35	77.8	388	2	S25298	extensin (clone To
398	35	77.8	190	2	T35570	hypothetical prote	471	35	77.8	388	2	G85147	hypothetical prote
399	35	77.8	192	2	T30477	hypothetical prote	472	35	77.8	389	2	D87460	ribonuclease D (im
400	35	77.8	199	2	T07622	extensin homolog -	473	35	77.8	391	2	S69192	serine O-acetyltra
401	35	77.8	209	2	C89005	protein T24A6.3 [1	474	35	77.8	392	2	B48423	homeotic protein e
402	35	77.8	214	2	A12371	hypothetical prote	475	35	77.8	395	2	I49575	CCAAT/enhancer bin
403	35	77.8	214	2	A47715	cardiac-specific h	476	35	77.8	396	2	S39793	MAPK-activated pro
404	35	77.8	216	2	T20186	hypothetical prote	477	35	77.8	398	2	JQ1245	hypothetical 43K p
405	35	77.8	218	2	F84748	probable AP2 domai	478	35	77.8	399	2	I49754	homeobox protein -
406	35	77.8	233	2	T17218	hypothetical prote	479	35	77.8	401	2	A48423	engrailed homeodom
407	35	77.8	249	2	T41847	AcMNPV orf106 - Bo	480	35	77.8	404	1	Q0BE13	BMRF1 protein - hu
408	35	77.8	250	2	D86214	hypothetical prote	481	35	77.8	408	2	G96707	hypothetical prote
409	35	77.8	250	2	T08071	L-ascorbate peroxi	482	35	77.8	414	2	T50010	hypothetical prote
410	35	77.8	250	2	S43157	L-ascorbate peroxi	483	35	77.8	416	2	G86232	cysteine proteinas
411	35	77.8	252	2	B48725	MDV specific prote	484	35	77.8	419	2	S56073	opaque-2 protein -
412	35	77.8	263	2	S20866	L-ascorbate peroxi	485	35	77.8	421	1	S11674	acrosin (EC 3.4.21
413	35	77.8	264	2	PQ0478	pistil extensin-li	486	35	77.8	423	2	T48121	hypothetical prote
414	35	77.8	264	2	T10572	hypothetical prote	487	35	77.8	427	2	B96804	muclain-like prot
415	35	77.8	265	2	T46834	hypothetical prote	488	35	77.8	432	2	T06782	extensin - soybean
416	35	77.8	265	2	T46089	proline-rich prote	489	35	77.8	432	2	A43448	thrombin receptor
417	35	77.8	267	2	S08314	cell wall glycopro	490	35	77.8	432	2	B96515	hypothetical prote
418	35	77.8	267	2	T15645	hypothetical prote	491	35	77.8	433	2	T07910	hydroxyproline-ric
419	35	77.8	268	2	H84684	En/Spm-like transp	492	35	77.8	434	2	B84684	hypothetical prote
420	35	77.8	275	2	T02559	probable spliceoso	493	35	77.8	436	2	T22253	phosphatidylcholin
421	35	77.8	280	2	T11671	extensin-like prot	494	35	77.8	438	1	XM5N	hypothetical prote
422	35	77.8	281	2	D70845	hypothetical prote	495	35	77.8	438	2	T12494	hypothetical prote
423	35	77.8	283	2	G01526	insulin promoter f	496	35	77.8	439	2	B86302	hypothetical prote
424	35	77.8	284	2	S39581	IPF1 protein - mou	497	35	77.8	440	2	I49681	glycerolaldehyde-3-p
425	35	77.8	289	2	A43562	homeotic protein H	498	35	77.8	440	2	S51614	Algal-CAM - Volvox
426	35	77.8	290	2	T07080	shock protein SRC2	499	35	77.8	442	2	A38592	retinoic acid rece
427	35	77.8	296	2	T43336	cah3 protein - fis	500	35	77.8	443	1	C35991	retinoic acid rece
428	35	77.8	300	2	T08453	hypothetical prote	501	35	77.8	443	2	S29324	transcription fact
429	35	77.8	302	2	A96661	unknown protein, 8	502	35	77.8	444	2	I51256	retinoic acid rece
430	35	77.8	303	2	S28264	hydroxyproline-ric	503	35	77.8	445	1	A49447	transcription fact
431	35	77.8	310	2	T45873	hypothetical prote	504	35	77.8	445	1	S31224	transcription fact
432	35	77.8	311	2	T03827	myb protein homolo	505	35	77.8	445	2	H96560	hypothetical prote
433	35	77.8	311	2	T02783	probable homeotic	506	35	77.8	446	2	A34418	H-2 region II bind
434	35	77.8	314	2	T48514	hypothetical prote	507	35	77.8	446	2	T07907	hydroxyproline-ric
435	35	77.8	318	2	T29479	hypothetical prote	508	35	77.8	447	2	B34714	retinoic acid rece
436	35	77.8	320	2	T18319	hypothetical prote	509	35	77.8	447	2	T49439	hypothetical prote
437	35	77.8	322	2	S25299	extensin precursor	510	35	77.8	448	2	D41727	retinoid X recepto
438	35	77.8	322	2	T22403	hypothetical prote	511	35	77.8	451	2	A41651	retinoic acid rece
439	35	77.8	325	2	T31474	hypothetical prote	512	35	77.8	454	1	A33903	retinoic acid rece
440	35	77.8	325	2	A55558	albumin D-box bind	513	35	77.8	454	2	S06124	retinoic acid rece
441	35	77.8	326	2	A45452	transcription fact	514	35	77.8	458	2	A34714	retinoic acid rece
442	35	77.8	328	2	JQ0985	probable lvtB prot	515	35	77.8	459	2	T26878	hypothetical prote
443	35	77.8	329	2	E70973	probable extensin	516	35	77.8	459	2	S03116	gene 33 protein, h
444	35	77.8	330	2	T05717	hypothetical prote	517	35	77.8	461	2	T10741	extensin-like prot
445	35	77.8	331	2	E96529	hypothetical prote	518	35	77.8	467	2	T34874	hypothetical prote
446	35	77.8	338	2	H96765	hypothetical prote	519	35	77.8	473	2	B85187	glycoprotein homol
447	35	77.8	341	2	T37502	hypothetical prote	520	35	77.8	476	2	T03170	hypothetical prote
448	35	77.8	345	2	T01348	hypothetical prote	521	35	77.8	476	2	B41977	retinoic acid rece
449	35	77.8	350	2	S22456	hydroxyproline-ric	522	35	77.8	477	2	T27051	hypothetical prote
450	35	77.8	350	2	G75571	MutT/nudix family	523	35	77.8	477	2	T46304	hypothetical prote
451	35	77.8	352	2	T33664	hypothetical prote	524	35	77.8	478	2	JC4940	synapsin IIb - hum
452	35	77.8	357	2	T47367	hypothetical prote	525	35	77.8	479	1	S15031	paired box transcr
453	35	77.8	358	2	JC4311	CCAAT/enhancer bin	526	35	77.8	485	2	A33647	sulfated surface g
454	35	77.8	358	2	A54265	CCAAT/enhancer-bin	527	35	77.8	485	2	PC4427	Nck, Ash and phosph
455	35	77.8	358	2	A75561	conserved hypotet	528	35	77.8	487	1	S52261	NADH2 dehydrogenas
456	35	77.8	365	2	S47657	peroxidase (EC 1.1	529	35	77.8	487	2	S42442	nuclear protein EB
457	35	77.8	365	2	A34894	hepatic transcript	530	35	77.8	489	2	T11622	extensin class 1 p
458	35	77.8	368	2	C29356	hydroxyproline-ric	531	35	77.8	494	2	B96534	hypothetical prote
459	35	77.8	368	2	T51200	hypothetical prote	532	35	77.8	499	2	I51257	retinoic acid rece
460	35	77.8	370	2	JC2204	MAPK-activated pro	533	35	77.8	502	2	A55197	Wiskott-Aldrich sy
461	35	77.8	375	2	S58484	gag protein - maiz	534	35	77.8	520	2	I84718	RXR-beta1 isoform
462	35	77.8	376	2	S45763	hypothetical prote	535	35	77.8	520	2	E97813	WASP, N-WASP, MENA
463	35	77.8	377	2	JC4368	protein farnesyltr	536	35	77.8	520	2	G86414	probable protein k
464	35	77.8	377	2	A41625	protein farnesyltr	537	35	77.8	520	2	S78502	paired box transcr
465	35	77.8	379	2	S42529	Opaque-2-related p	538	35	77.8	521	2	S54266	Glycoprotein GC -
466	35	77.8	379	2	S31719	proline-rich prote	539	35	77.8	529	1	W7AD22	early E2A DNA-bind
467	35	77.8	381	2	T27806	hypothetical prote	540	35	77.8	529	1	W7AD25	early E2A DNA-bind

541	35	77.8	530	2	A45690	transactivator EBN	614	35	77.8	981	2	A41401	mineralocorticoid
542	35	77.8	532	2	G84775	probable E2F5 fami	615	35	77.8	984	2	A29513	mineralocorticoid
543	35	77.8	533	2	S37781	retinoid x recepto	616	35	77.8	1001	2	T28897	hypothetical prote
544	35	77.8	534	1	A48529	ubiquinol-cytochro	617	35	77.8	1015	2	T42013	frequency clock pr
545	35	77.8	540	2	S44830	F54F2.5 protein -	618	35	77.8	1032	2	D83637	serine/threonine p
546	35	77.8	542	2	A44358	zyxin - chicken	619	35	77.8	1039	2	T22117	hypothetical prote
547	35	77.8	543	1	ERADDG	fiber protein - ca	620	35	77.8	1040	2	T29092	TSC-22 protein hom
548	35	77.8	548	2	S52735	CW17R protein - mo	621	35	77.8	1047	2	A55617	masquerade precurs
549	35	77.8	550	2	G70597	probable proteinas	622	35	77.8	1091	2	T13170	diaphanous protein
550	35	77.8	554	2	S75969	hypothetical prote	623	35	77.8	1092	2	T18305	replication factor
551	35	77.8	557	2	S62522	nuclear protein SP	624	35	77.8	1092	2	T18306	replication factor
552	35	77.8	561	2	T23722	hypothetical prote	625	35	77.8	1100	2	T30967	transcription acti
553	35	77.8	568	2	JC7317	cyclin-dependent k	626	35	77.8	1118	2	A48292	mucin, tracheobron
554	35	77.8	571	2	G84426	hypothetical prote	627	35	77.8	1119	2	T16720	hypothetical prote
555	35	77.8	572	2	G02845	zyxin - human	628	35	77.8	1173	2	T31421	C-terminal domain-
556	35	77.8	575	2	G88346	protein F42G4.3a [629	35	77.8	1176	2	T49482	hypothetical prote
557	35	77.8	580	2	T10863	extensin precursor	630	35	77.8	1188	2	S49915	extensin-like prot
558	35	77.8	581	2	E86408	F3H9.11 protein -	631	35	77.8	1209	2	T52523	hypothetical prote
559	35	77.8	585	2	I58403	H4 protein - human	632	35	77.8	1213	2	A41724	limb deformity (ld
560	35	77.8	599	2	T10798	pheraphorin-S - Vo	633	35	77.8	1252	2	T14272	cortactin-binding
561	35	77.8	603	2	T22111	hypothetical prote	634	35	77.8	1276	2	E96776	hypothetical prote
562	35	77.8	605	1	Q0BE3R	BVRF2 (EC-RF3) pro	635	35	77.8	1296	2	T13936	collar protein iso
563	35	77.8	620	2	S06733	hydroxyproline-ric	636	35	77.8	1297	2	S25714	son-of-sevenless-2
564	35	77.8	633	2	F84564	probable protein k	637	35	77.8	1306	2	T13592	hypothetical prote
565	35	77.8	639	2	T13151	adapter protein CM	638	35	77.8	1366	2	B84924	hypothetical prote
566	35	77.8	639	2	G02919	transcription fact	639	35	77.8	1390	2	T31353	polyprotein - Arab
567	35	77.8	639	2	H86382	hypothetical prote	640	35	77.8	1392	2	T51947	probable transcrip
568	35	77.8	644	2	S15464	gp70 protein - mur	641	35	77.8	1400	2	T52359	hypothetical prote
569	35	77.8	653	2	JC8013	Ran-binding protei	642	35	77.8	1429	2	T13720	gene expanded prot
570	35	77.8	658	2	T08153	cysteine proteinas	643	35	77.8	1584	2	T00026	brain-specific ang
571	35	77.8	664	2	C84747	probable protein k	644	35	77.8	1621	2	T15264	hypothetical prote
572	35	77.8	664	2	T01368	hypothetical prote	645	35	77.8	1736	2	A47747	tight junction pro
573	35	77.8	666	2	T31461	probable magnesium	646	35	77.8	1878	2	B86189	hypothetical prote
574	35	77.8	678	2	AF1841	Mg chelatase chain	647	35	77.8	1905	2	T18267	multidrug resistan
575	35	77.8	694	2	T01005	hypothetical prote	648	35	77.8	1952	2	T48814	hypothetical prote
576	35	77.8	700	2	S09699	bib protein - frui	649	35	77.8	1953	2	S63244	BNI1 protein - yea
577	35	77.8	707	2	S60588	drebrin A - rat	650	35	77.8	2109	2	T17490	polyketide synthas
578	35	77.8	711	2	S43464	ecdysteroid-induce	651	35	77.8	2363	2	T38841	probable pre-mRNA
579	35	77.8	716	2	G01627	androgen receptor	652	35	77.8	2500	2	G88493	protein F5789.2 [i
580	35	77.8	731	2	B86369	hypothetical prote	653	35	77.8	2555	2	A40043	notch protein homo
581	35	77.8	734	2	T04876	hypothetical prote	654	35	77.8	2606	2	T03159	large tegument pro
582	35	77.8	736	2	JH0681	gephyrin - rat	655	35	77.8	2706	2	T28155	variant-specific s
583	35	77.8	737	2	S28030	DNA-binding protei	656	35	77.8	2810	2	T22298	hypothetical prote
584	35	77.8	744	1	Q0BEA7	UL69 protein - hum	657	35	77.8	3119	2	T49729	HD protein - mouse
585	35	77.8	744	2	E86235	hypothetical prote	658	35	77.8	3164	1	WMBEH6	U336 protein - hum
586	35	77.8	745	2	T51370	hypothetical prote	659	34.5	76.7	407	2	T24951	hypothetical prote
587	35	77.8	754	2	T06249	protoporphyrin IX	660	34.5	76.7	409	1	VCW2S	env polyprotein (v
588	35	77.8	755	2	T47731	hypothetical prote	661	34.5	76.7	437	2	B82778	cell cycle protein
589	35	77.8	758	2	S60586	glucocorticoid rec	662	34.5	76.7	886	2	A54442	3',5'-cyclic-nucle
590	35	77.8	758	2	T02925	protoporphyrin IX	663	34	75.6	10	2	A36454	trypsin-modulating
591	35	77.8	760	2	B96724	hypothetical prote	664	34	75.6	15	2	PT0037	light harvesting c
592	35	77.8	762	2	C96653	hypothetical prote	665	34	75.6	52	2	F71353	hypothetical prote
593	35	77.8	764	2	A56208	DNA-binding protei	666	34	75.6	57	2	S10782	salivary protein p
594	35	77.8	772	2	T13078	KIAA0992 protein -	667	34	75.6	65	2	T06457	outer envelope mem
595	35	77.8	778	2	B86218	protein T2767.20 [668	34	75.6	66	2	T31844	hypothetical prote
596	35	77.8	782	2	A82940	hypothetical prote	669	34	75.6	74	2	B40513	hypothetical prote
597	35	77.8	786	2	T01456	extensin homolog F	670	34	75.6	74	2	T33086	hypothetical prote
598	35	77.8	789	2	T52067	hypothetical prote	671	34	75.6	76	2	C38355	basic proline-rich
599	35	77.8	791	2	C82940	hypothetical prote	672	34	75.6	77	2	T15349	hypothetical prote
600	35	77.8	805	2	T41810	AcMPV orf66 - Bom	673	34	75.6	79	1	PJHUB	proline-rich pepti
601	35	77.8	808	2	C72858	AcOrf-66 protein -	674	34	75.6	79	2	E87607	hypothetical prote
602	35	77.8	811	2	S04085	ovarian tumor prot	675	34	75.6	81	2	T30482	zink finger protei
603	35	77.8	817	2	T03852	protein phosphatas	676	34	75.6	90	2	T18068	hypothetical prote
604	35	77.8	820	2	T00645	hypothetical prote	677	34	75.6	91	2	T11554	vxp protein - simi
605	35	77.8	845	2	T17291	hypothetical prote	678	34	75.6	91	2	T30512	hypothetical prote
606	35	77.8	846	2	T21700	hypothetical prote	679	34	75.6	92	2	T34146	hypothetical prote
607	35	77.8	847	2	F96531	hypothetical prote	680	34	75.6	103	2	T29597	hypothetical prote
608	35	77.8	884	2	D96730	unknown protein F5	681	34	75.6	104	2	G84607	hypothetical prote
609	35	77.8	899	2	A35895	androgen receptor	682	34	75.6	105	2	T22564	hypothetical prote
610	35	77.8	902	2	B40494	androgen receptor	683	34	75.6	106	2	B72490	hypothetical prote
611	35	77.8	958	2	T13593	hypothetical prote	684	34	75.6	108	2	T26880	hypothetical prote
612	35	77.8	962	2	T00262	hypothetical prote	685	34	75.6	108	2	A48831	vitelline membrane
613	35	77.8	975	2	T48107	hypothetical prote	686	34	75.6	108	2	F72653	hypothetical prote

687	34	75.6	111	2	S08438	vpv protein - huma	760	34	75.6	188	2	C96593	unknown protein, 9
688	34	75.6	112	1	ASLQGH	vpv protein - huma	761	34	75.6	189	2	T10580	zinc-finger protei
689	34	75.6	112	1	ASLQJX	vpv protein - huma	762	34	75.6	189	2	E90769	EspF-like protein
690	34	75.6	112	2	T11562	vpv protein - simi	763	34	75.6	191	2	T28682	hypothetical prote
691	34	75.6	112	2	S07990	vpv protein - simi	764	34	75.6	194	2	A38203	proline-rich prote
692	34	75.6	112	2	S57448	DNA binding protei	765	34	75.6	195	2	T07735	modulin-20a - soyb
693	34	75.6	115	2	T36886	hypothetical prote	766	34	75.6	195	2	T21300	hypothetical prote
694	34	75.6	117	2	T46322	hypothetical prote	767	34	75.6	195	2	I52578	ALL-1 protein - hu
695	34	75.6	121	2	S17718	anther-specific pr	768	34	75.6	196	2	B44525	Ca2+-transporting
696	34	75.6	121	2	S12245	anther-specific pr	769	34	75.6	197	2	S35252	proline-rich prote
697	34	75.6	129	2	T19263	hypothetical prote	770	34	75.6	198	2	E86261	R13K23.6 protein -
698	34	75.6	133	2	T26755	hypothetical prote	771	34	75.6	202	2	T11744	dehydrin - kidney
699	34	75.6	133	2	H87681	hypothetical prote	772	34	75.6	204	2	T07679	protein import rec
700	34	75.6	134	2	D84672	hypothetical prote	773	34	75.6	205	2	C44525	Ca2+-transporting
701	34	75.6	134	2	T36365	proline-rich prote	774	34	75.6	205	2	F87623	hypothetical prote
702	34	75.6	135	2	E96750	hypothetical prote	775	34	75.6	207	2	I53154	scleraxis - mouse
703	34	75.6	137	2	T22308	hypothetical prote	776	34	75.6	209	2	T07779	dehydrin homolog C
704	34	75.6	139	2	T37932	very hypothetical	777	34	75.6	211	2	B89716	protein F4588.3 [i
705	34	75.6	139	2	A71099	hypothetical prote	778	34	75.6	212	2	S74288	hypothetical prote
706	34	75.6	139	2	C87459	hypothetical prote	779	34	75.6	213	2	T48490	embryo-specific pr
707	34	75.6	139	2	S61985	extensin precursor	780	34	75.6	214	2	S65052	pistil-specific pr
708	34	75.6	139	4	S09612	hypothetical xrepB	781	34	75.6	215	2	T22572	hypothetical prote
709	34	75.6	141	2	T06646	blue copper-bindin	782	34	75.6	217	2	T09965	extensin CVC17 pr
710	34	75.6	141	2	A34043	hypothetical proli	783	34	75.6	217	2	T15873	hypothetical prote
711	34	75.6	142	2	T51039	related to extensi	784	34	75.6	217	4	I65366	ALL-1/AF-4 mutat
712	34	75.6	143	2	S42579	QID3 protein - fun	785	34	75.6	218	2	T23318	hypothetical prote
713	34	75.6	143	2	E72699	hypothetical prote	786	34	75.6	220	2	S52005	modulin-30 (Npv30)
714	34	75.6	145	2	E88462	protein C05H8.2 [i	787	34	75.6	220	2	S42879	modulin-30 - kidne
715	34	75.6	145	2	T48552	glutaredoxin-like	788	34	75.6	221	2	H82857	TonB protein XF000
716	34	75.6	148	2	S46956	outer membrane pro	789	34	75.6	222	2	T43500	hypothetical prote
717	34	75.6	148	2	S46542	outer envelope pro	790	34	75.6	223	2	A42817	proline-rich prote
718	34	75.6	150	2	C72697	hypothetical prote	791	34	75.6	224	2	G87377	hypothetical prote
719	34	75.6	151	2	S10084	hypothetical 16K p	792	34	75.6	225	2	T09964	extensin CVC15 pre
720	34	75.6	153	2	T31654	hypothetical prote	793	34	75.6	225	2	D88492	protein T07E3.6 [i
721	34	75.6	154	2	T41831	AcMNPV orf51 - Bom	794	34	75.6	226	2	S41032	hypothetical prote
722	34	75.6	155	2	T16855	hypothetical prote	795	34	75.6	226	2	T35236	hypothetical prote
723	34	75.6	156	2	T29730	hypothetical prote	796	34	75.6	230	2	T36672	membrane-spanning
724	34	75.6	157	2	T15064	protein F6N18.7 [i	797	34	75.6	231	2	T50853	response regulator
725	34	75.6	158	2	A86452	hypothetical prote	798	34	75.6	231	2	S37108	cuticlin 2 - Caeno
726	34	75.6	159	2	D72486	hypothetical prote	799	34	75.6	235	2	E75342	conserved hypothet
727	34	75.6	160	2	S58759	eazrin - rat (fragm	800	34	75.6	236	2	T45835	hypothetical prote
728	34	75.6	160	2	AC2958	conserved hypothet	801	34	75.6	236	2	T02577	probable Ap2 domai
729	34	75.6	161	2	T45055	hypothetical prote	802	34	75.6	239	2	T03078	conserved hypothet
730	34	75.6	163	2	D84674	hypothetical prote	803	34	75.6	240	2	T25814	hypothetical prote
731	34	75.6	163	2	T33130	hypothetical prote	804	34	75.6	240	2	D70894	probable pra prote
732	34	75.6	164	2	C98325	probable transcrip	805	34	75.6	241	2	T22216	hypothetical prote
733	34	75.6	165	2	C71717	hypothetical prote	806	34	75.6	243	2	S28444	tonB protein - Pse
734	34	75.6	165	2	T24470	hypothetical prote	807	34	75.6	244	2	B87417	hypothetical prote
735	34	75.6	166	1	PIHUSC	salivary proline-r	808	34	75.6	244	2	S44822	F44E2.3 protein -
736	34	75.6	166	2	B25372	salivary proline-r	809	34	75.6	246	2	T46446	hypothetical prote
737	34	75.6	167	2	A44525	Ca2+-transporting	810	34	75.6	247	2	T06721	hypothetical prote
738	34	75.6	169	2	T18321	hypothetical prote	811	34	75.6	247	2	T50874	hypothetical cytoc
739	34	75.6	171	2	A27307	proline-rich phosp	812	34	75.6	247	2	T33469	hypothetical prote
740	34	75.6	172	2	S68232	antimicrobial prot	813	34	75.6	248	2	F83558	hypothetical prote
741	34	75.6	172	2	D29149	proline-rich prote	814	34	75.6	249	2	A41497	36K antigen pra -
742	34	75.6	172	2	B41132	collagen-related p	815	34	75.6	250	1	A31757	homeotic protein H
743	34	75.6	172	2	T27505	hypothetical prote	816	34	75.6	250	2	B81734	conserved hypothet
744	34	75.6	173	2	T51469	hypothetical prote	817	34	75.6	250	2	T08908	hypothetical prote
745	34	75.6	174	2	A41358	glycine/proline-ri	818	34	75.6	250	2	T16342	hypothetical prote
746	34	75.6	174	2	B96543	gag polyprotein -	819	34	75.6	250	2	A85632	hypothetical prote
747	34	75.6	175	2	S71560	early light-induce	820	34	75.6	251	1	B60492	homeotic protein H
748	34	75.6	180	2	E84768	hypothetical prote	821	34	75.6	253	2	T32465	hypothetical prote
749	34	75.6	180	2	D72629	hypothetical prote	822	34	75.6	254	2	S52236	MADS box protein a
750	34	75.6	181	2	JC5233	spermatopharin Sp2	823	34	75.6	254	2	B84901	hypothetical prote
751	34	75.6	181	2	T48558	hypothetical prote	824	34	75.6	254	2	AE2516	hypothetical prote
752	34	75.6	182	2	A44157	spermatopharin Sp2	825	34	75.6	255	2	C36222	phosphoprotein pho
753	34	75.6	182	2	B35650	hypothetical 20K p	826	34	75.6	255	2	T34163	hypothetical prote
754	34	75.6	183	2	JC4356	ankyrin precursor	827	34	75.6	256	2	S39794	MAPK-activated pro
755	34	75.6	184	2	S78091	endocuticular prot	828	34	75.6	262	2	T23241	hypothetical prote
756	34	75.6	185	2	B26669	modulin-20 precurs	829	34	75.6	263	2	T12698	extensin - common
757	34	75.6	185	2	B84531	hypothetical prote	830	34	75.6	264	2	B85005	hypothetical prote
758	34	75.6	186	2	C41132	collagen-related p	831	34	75.6	268	2	T70823	hypothetical prote
759	34	75.6	187	2	S38036	hypothetical prote	832	34	75.6	269	2	C84707	hypothetical prote

833	34	75.6	269	2	T30468	hypothetical prote	906	34	75.6	350	2	T14191	extensin homolog T
834	34	75.6	269	2	A71401	hypothetical prote	907	34	75.6	352	2	C96643	hypothetical prote
835	34	75.6	270	2	A84549	probable C2H2-type	908	34	75.6	354	1	GNVSR	genome polyprotein
836	34	75.6	270	2	S63049	hypothetical prote	909	34	75.6	355	2	S41285	coat protein - swe
837	34	75.6	271	2	T26640	hypothetical prote	910	34	75.6	356	1	WJHU2H	homeotic protein H
838	34	75.6	271	2	T49956	hypothetical prote	911	34	75.6	357	2	PC4293	nuclear factor 1 f
839	34	75.6	273	2	F87414	phage spo1 DNA pol	912	34	75.6	358	2	T10244	G-Box binding prot
840	34	75.6	274	2	T04619	hypothetical prote	913	34	75.6	358	2	E87309	hypothetical prote
841	34	75.6	275	2	S75818	hypothetical prote	914	34	75.6	358	2	T01296	leucine-rich repea
842	34	75.6	277	2	H87552	hypothetical prote	915	34	75.6	359	2	T13478	hypothetical prote
843	34	75.6	277	2	S78063	homeobox protein P	916	34	75.6	360	2	D87332	hypothetical prote
844	34	75.6	279	2	T02495	hypothetical prote	917	34	75.6	364	2	I48188	gene NK6.1 protei
845	34	75.6	280	2	A53027	transcription fact	918	34	75.6	365	2	T04718	hypothetical prote
846	34	75.6	281	2	T29150	hypothetical prote	919	34	75.6	365	2	T24955	hypothetical prote
847	34	75.6	281	2	S36748	transcription fact	920	34	75.6	365	2	E42832	factor VIII intron
848	34	75.6	282	2	T26635	hypothetical prote	921	34	75.6	366	2	T30457	hypothetical prote
849	34	75.6	282	2	A53336	transcription fact	922	34	75.6	366	2	I53035	trithorax homolog
850	34	75.6	283	2	S13383	hydroxyproline-ric	923	34	75.6	367	1	S02193	cellular tumor ant
851	34	75.6	287	2	T34397	hypothetical prote	924	34	75.6	368	2	T03828	myb protein - rice
852	34	75.6	288	2	T45715	hypothetical prote	925	34	75.6	370	2	B88455	protein T15812.2 (
853	34	75.6	289	2	T52354	hypothetical prote	926	34	75.6	370	2	JC6130	paired box transcr
854	34	75.6	289	2	A62006	hypothetical prote	927	34	75.6	372	2	S74859	hypothetical prote
855	34	75.6	290	2	AC2030	hypothetical prote	928	34	75.6	372	2	T10472	G-box binding prot
856	34	75.6	291	2	T27534	hypothetical prote	929	34	75.6	373	2	E84595	hypothetical prote
857	34	75.6	294	2	A55477	survival motor neu	930	34	75.6	373	2	A70856	probable lppz prot
858	34	75.6	297	2	B83049	conserved hypothet	931	34	75.6	376	2	A26066	segmentation prote
859	34	75.6	301	2	T18788	hypothetical prote	932	34	75.6	378	2	T21651	hypothetical prote
860	34	75.6	302	2	T50001	hypothetical prote	933	34	75.6	379	2	S49999	hypothetical prote
861	34	75.6	302	2	S13865	chlorophyll a/b-bi	934	34	75.6	381	2	G02668	neurogenic basic-h
862	34	75.6	303	1	TVH9UD	transforming prote	935	34	75.6	381	2	S59679	hypothetical prote
863	34	75.6	304	2	A32993	transcription fact	936	34	75.6	381	2	B86175	protein P19P19.21
864	34	75.6	304	2	T48281	hypothetical prote	937	34	75.6	382	2	E85082	hypothetical prote
865	34	75.6	304	2	H86332	T20H2.26 protein -	938	34	75.6	382	2	T14186	hypothetical prote
866	34	75.6	308	2	A86246	hypothetical prote	939	34	75.6	383	2	T00674	hypothetical prote
867	34	75.6	309	1	S09257	homeotic protein H	940	34	75.6	384	2	H88924	protein F33H1.2 (
868	34	75.6	309	2	E36596	nuclear factor I (941	34	75.6	384	2	S51796	vasodilator-stimul
869	34	75.6	309	2	T47685	probable RNA bindi	942	34	75.6	384	2	C85816	unknown protein en
870	34	75.6	311	2	S20085	transcription fact	943	34	75.6	385	2	T20410	hypothetical prote
871	34	75.6	312	2	D87475	rare lipoprotein A	944	34	75.6	388	2	T26922	hypothetical prote
872	34	75.6	312	2	T29232	hypothetical prote	945	34	75.6	389	2	C70921	probable molybdopt
873	34	75.6	313	1	FOVDA	gag polyprotein -	946	34	75.6	389	2	T30411	protein RFL - rice
874	34	75.6	314	2	D71410	hypothetical prote	947	34	75.6	389	2	S27200	proline-rich prote
875	34	75.6	315	2	JQ1250	coat protein - chr	948	34	75.6	389	2	B82819	outer membrane pro
876	34	75.6	315	2	T34332	hypothetical prote	949	34	75.6	390	2	G01936	Abl binding protei
877	34	75.6	315	2	A64689	hypothetical prote	950	34	75.6	392	2	F84650	probable protein p
878	34	75.6	315	2	A56561	35K proline-rich p	951	34	75.6	393	2	T33103	lin-1 protein - Ca
879	34	75.6	323	2	S20099	transforming prote	952	34	75.6	393	2	JC5614	RNA6 protein - rat
880	34	75.6	323	2	H70765	hypothetical prote	953	34	75.6	394	2	C84905	probable extensin
881	34	75.6	324	2	S13497	cAMP-binding prote	954	34	75.6	396	1	A58938	surface protein rh
882	34	75.6	324	2	T24102	hypothetical prote	955	34	75.6	396	2	T07021	extensin-like prot
883	34	75.6	328	2	C42756	hypothetical prote	956	34	75.6	396	2	A36339	F1O protein - gard
884	34	75.6	329	2	T10064	cytokinin-induced	957	34	75.6	396	2	S41224	hnRNP protein - Af
885	34	75.6	330	2	S37595	mucin JUI10 - huma	958	34	75.6	396	2	S62964	hypothetical prote
886	34	75.6	331	2	B47236	zinc-finger protei	959	34	75.6	396	2	S75266	hypothetical prote
887	34	75.6	333	2	S11484	CABP1 protein - sl	960	34	75.6	398	2	A13231	conjugal transfer
888	34	75.6	334	2	T04572	hypothetical prote	961	34	75.6	399	2	T01345	hypothetical prote
889	34	75.6	337	2	C90968	Egfp-like protein	962	34	75.6	404	2	H88650	protein C09G12.9 (
890	34	75.6	338	1	TVMSFB	transforming prote	963	34	75.6	405	2	T42663	hypothetical prote
891	34	75.6	338	2	I53043	transforming prote	964	34	75.6	406	2	T36632	probable oxidoredu
892	34	75.6	338	2	I56893	transcription fact	965	34	75.6	407	1	A40150	inhibin beta-B cha
893	34	75.6	340	2	S53763	homeotic protein H	966	34	75.6	410	2	C70889	hypothetical prote
894	34	75.6	344	1	AHMS	lg alpha chain C r	967	34	75.6	411	2	B41398	inhibin beta-B cha
895	34	75.6	344	2	A86294	hypothetical prote	968	34	75.6	413	2	T03240	FLO/LFY protein ho
896	34	75.6	345	2	S12788	transcription fact	969	34	75.6	413	2	H87604	hypothetical prote
897	34	75.6	345	2	A48462	dense granule prot	970	34	75.6	414	2	C86301	hypothetical prote
898	34	75.6	345	2	C86272	F7A19.1 protein -	971	34	75.6	415	1	A34170	arginine/serine-ri
899	34	75.6	346	2	S19129	proline-rich prote	972	34	75.6	416	2	T03243	acrosin (EC 3.4.21
900	34	75.6	346	2	S76923	hypothetical prote	973	34	75.6	417	2	T15174	FLO/LFY protein ho
901	34	75.6	347	2	A43815	transforming prote	974	34	75.6	418	2	F70940	hypothetical prote
902	34	75.6	348	2	A49594	enhancer factor pr	975	34	75.6	418	2	T16713	hypothetical prote
903	34	75.6	348	2	AB3260	hypothetical membr	976	34	75.6	419	2	T19871	hypothetical prote
904	34	75.6	348	2	S52720	homeobox protein g	977	34	75.6	420	2	T46910	hypothetical prote
905	34	75.6	349	2	G84852	hypothetical prote	978	34	75.6	420	2	JC4716	zinc finger DNA-bi

979 34 75.6 427 2 T48159
980 34 75.6 427 2 I49603
981 34 75.6 428 1 TVHUEK
982 34 75.6 428 2 E71415
983 34 75.6 429 2 T06296
984 34 75.6 429 2 A47305
985 34 75.6 430 2 JC2301
986 34 75.6 431 2 S47538
987 34 75.6 431 2 T12450
988 34 75.6 433 1 FOLJH2
989 34 75.6 434 2 C96315
990 34 75.6 435 2 A42672
991 34 75.6 436 2 B55452
992 34 75.6 439 2 AF1048
993 34 75.6 440 2 S48953
994 34 75.6 440 2 H86312
995 34 75.6 440 2 JC7807
996 34 75.6 440 2 S71795
997 34 75.6 441 1 TVCHMC
998 34 75.6 441 2 T22531
999 34 75.6 442 2 T34018
1000 34 75.6 443 2 G86388

ALIGNMENTS

RESULT 1
A53692
synapsin I - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
C;Accession: A53692
R;Chin, L.S.; Li, L.; Greengard, P.
J. Biol. Chem. 269, 18507-18513, 1994
A;Title: Neuron-specific expression of the synapsin II gene is directed by a specific co
A;Reference number: A53692; MUID:94308086; PMID:8034599
A;Accession: A53692
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-125 <RES>
A;Cross-references: UNIPROT:O88935; UNIPARC:UPI0000016D062; GB:L32025; NID:g1041084; PIDN:
C;Genetics:
A;Gene: SYN1

Query Match 100.0%; Score 45; DB 2; Length 125;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPOPPPP 7
|||
DB 27 RPOPPPP 33

RESULT 2
A25704
synapsin I - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 21-May-1988 #sequence_revision 21-May-1988 #text_change 09-Jul-2004
C;Accession: A25704
R;McCaffery, C.A.; DeGennaro, L.J.
EMBO J. 5, 3167-3173, 1986
A;Title: Determination and analysis of the primary structure of the nerve terminal speci
A;Reference number: A25704; MUID:87133474; PMID:3028773
A;Accession: A25704
A;Molecule type: mRNA
A;Residues: 1-691 <MCC>
A;Cross-references: UNIPROT:P09951; UNIPARC:UPI00000170B12; GB:X04655; NID:g57181; PIDN:
A;Experimental source: brain
C;Keywords: actin binding; alternative splicing; phosphoprotein
Query Match 100.0%; Score 45; DB 2; Length 691;
Best Local Similarity 100.0%; Pred. No. 60;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RPOPPPP 7
|||
DB 14 RPOPPPP 20

RESULT 3

A30411
synapsin Ia - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 23-Oct-1990 #sequence_revision 23-Oct-1990 #text_change 09-Jul-2004
C;Accession: A30411; B30411
R;Suedhof, T.C.; Czernik, A.J.; Kao, H.T.; Takei, K.; Johnston, P.A.; Horiuchi, A.; Kanai
Science 245, 1474-1480, 1989
A;Title: Synapsins: mosaics of shared and individual domains in a family of synaptic vesi
A;Reference number: A30411; MUID:89388265; PMID:2506642
A;Accession: A30411
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-704 <SUE>
A;Cross-references: UNIPROT:P09951; UNIPARC:UPI000004368E; GB:M27812; NID:g206920; PIDN:
A;Accession: B30411
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-659; 'KA', 662; 'PQAQAP', <SU2>
A;Cross-references: UNIPARC:UPI000002B48E; GB:M27924; NID:g206932; PIDN:AAA42148.1; PID:
C;Keywords: actin binding; alternative splicing; phosphoprotein

Query Match 100.0%; Score 45; DB 2; Length 704;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPOPPPP 7
|||
DB 27 RPOPPPP 33

RESULT 4

A53363
synapsin I splice form a - human
C;Species: Homo sapiens (man)
C;Date: 17-Aug-1990 #sequence_revision 17-Aug-1990 #text_change 09-Jul-2004
C;Accession: A53363; B35363; A35805
R;Suedhof, T.C.
J. Biol. Chem. 265, 7849-7852, 1990
A;Title: The structure of the human synapsin I gene and protein.
A;Reference number: A53363; MUID:90243651; PMID:2110562
A;Accession: A53363
A;Molecule type: DNA
A;Residues: 1-705 <SUE>
A;Cross-references: UNIPROT:P17600; UNIPARC:UPI0000179AF3; GB:M58371; GB:J05431
A;Status: preliminary
A;Accession: B35363
A;Molecule type: DNA
A;Residues: 1-659; 'KASPAQAP', <SU2>
A;Cross-references: UNIPARC:UPI0000179AF4; GB:M58378; GB:J05431
R;Sauerwald, A.; Hoesche, C.; Oschwald, R.; Kilimann, M.W.
J. Biol. Chem. 265, 14932-14937, 1990
A;Title: The 5'-flanking region of the synapsin I gene. A G+C-rich, TATA- and CAAT-less,
A;Reference number: A35805; MUID:90368667; PMID:2118519
A;Accession: A35805
A;Molecule type: DNA
A;Residues: 1-125 <SAU>
A;Cross-references: UNIPARC:UPI000016B07E; GB:M55301; NID:g338655; PIDN:AAA60608.1; PID:
C;Genetics:
A;Gene: GDB:SYN1
A;Cross-references: GDB:119606; OMIM:313440
A;Map position: Xp11.23-Xp11.23
C;Keywords: actin binding; alternative splicing; phosphoprotein

Query Match 100.0%; Score 45; DB 2; Length 705;
Best Local Similarity 100.0%; Pred. No. 61;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
|||||
DB 27 RPQPPPP 33

RESULT 5
E30411
synapsin Ia - bovine
C;Species: Bos primigenius taurus (cattle)
C;Date: 23-Oct-1990 #sequence_revision 23-Oct-1990 #text_change 01-Dec-2000
C;Accession: E30411; F30411; A35758
R;Suedhof, T.C.; Czernik, A.J.; Kao, H.T.; Takei, K.; Johnston, P.A.; Horiuchi, A.; Kana
Science 245, 1474-1480, 1989
A;Title: Synapsins: mosaics of shared and individual domains in a family of synaptic ves
A;Reference number: A30411; MUID:89388265; PMID:2506642
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-706 <SUE>
A;Cross-references: UNIPARC:UPI0000179AF5
A;Accession: F30411
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-661, 'KA', 664, 'PQAQAP' <SU2>
A;Cross-references: UNIPARC:UPI0000179AF5
R;Hall, F.L.; Mitchell, J.P.; Vulliamt, P.R.
J. Biol. Chem. 265, 6944-6948, 1990
A;Title: Phosphorylation of synapsin I at a novel site by proline-directed protein kinase
A;Reference number: A35758; MUID:90216728; PMID:2108963
A;Accession: A35758
A;Molecule type: protein
A;Residues: 532-556 <HAL>
A;Cross-references: UNIPARC:UPI0000179AF7
C;Keywords: actin binding; alternative splicing; phosphoprotein
F;551/Binding site: phosphate (Ser) (covalent) (by proline-directed kinase) #status expe

Query Match 100.0%; Score 45; DB 2; Length 706;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
|||||
DB 27 RPQPPPP 33

RESULT 6
T03455
ALR protein - human
C;Species: Homo sapiens (man)
C;Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 09-Jul-2004
C;Accession: T03455
R;Prasad, R.; Zhadanov, A.B.; Sedkov, Y.; Bullrich, F.; Druck, T.; Rallapalli, R.; Yano,
Oncogene 15, 549-560, 1997
A;Title: Structure and expression pattern of human ALR, a novel gene with strong homolog
A;Reference number: Z14954; MUID:97388474; PMID:9247308
A;Accession: T03455
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-4957 <PRA>
A;Cross-references: UNIPROT:O14686; UNIPARC:UPI00001100F1; EMBL:AF010404; NID:G2358286;
C;Genetics:
A;Gene: ALR
A;Map position: 12
C;Superfamily: acute lymphoblastic leukemia protein, ALR type
C;Keywords: alternative splicing

Query Match 100.0%; Score 45; DB 2; Length 4957;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7

DB 1806 RPQPPPP 1812
|||||

RESULT 7
T03454
ALR protein - human
C;Species: Homo sapiens (man)
C;Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 09-Jul-2004
C;Accession: T03454
R;Prasad, R.; Zhadanov, A.B.; Sedkov, Y.; Bullrich, F.; Druck, T.; Rallapalli, R.; Yano,
Oncogene 15, 549-560, 1997
A;Title: Structure and expression pattern of human ALR, a novel gene with strong homolog
A;Reference number: Z14954; MUID:97388474; PMID:9247308
A;Accession: T03454
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-5262 <PRA>
A;Cross-references: UNIPROT:O14686; UNIPARC:UPI0000110107; EMBL:AF010403; NID:G2358284;
C;Genetics:
A;Gene: ALR
A;Map position: 12
C;Superfamily: acute lymphoblastic leukemia protein, ALR type
C;Keywords: alternative splicing

Query Match 100.0%; Score 45; DB 2; Length 5262;
Best Local Similarity 100.0%; Pred. No. 4.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
|||||
DB 2111 RPQPPPP 2117

RESULT 8
T05847
hypothetical protein F17L22.180 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C;Accession: T05847
R;Bevan, M.; Van Der Schueren, J.; Chuang, Y.J.; Voet, M.; Robben, J.; Voickaert, G.; Bar
submitted to the Protein Sequence Database, February 1999
A;Reference number: Z15454
A;Accession: T05847
A;Molecule type: DNA
A;Residues: 1-139 <BEV>
A;Cross-references: UNIPROT:Q9SVS5; UNIPARC:UPI00000AA5A4; EMBL:AL035527
A;Experimental source: cultivar Columbia; BAC clone F17L22
C;Genetics:
A;Map position: 4
A;Introns: 26/1; 44/1; 78/3; 91/3
A;Note: F17L22.180

Query Match 93.3%; Score 42; DB 2; Length 139;
Best Local Similarity 85.7%; Pred. No. 31;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
|||||
DB 113 KPQPPPP 119

RESULT 9
AE2842
conserved hypothetical protein Atu2166 [imported] - Agrobacterium tumefaciens (strain C58
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C;Accession: AE2842
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClellan
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, E

```

ster, B.W.
A/Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A/Reference number: AB2577; MUID:21608550; PMID:11743193
A/Accession: AE2842
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-320 <KUR>
A/Cross-references: UNIPROT:Q8UDGO; UNIPARC:UPI00001646C6; GB:AE008688; PIDN:AAL43155.1;
A/Experimental source: strain C58 (Dupont)
C/Genetics:
A/Gene: Atu2166
A/Map position: circular chromosome

Query Match 93.3%; Score 42; DB 2; Length 320;
Best Local Similarity 85.7%; Pred. No. 69;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPOPPPPP 7
Db 74 KPQPPPPP 80
:|||||

RESULT 10
F97619
hypothetical protein AGR_C_3928 [imported] - Agrobacterium tumefaciens (strain C58, Cere
C/Species: Agrobacterium tumefaciens
C/Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C/Accession: F97619
R/Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A/Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A/Reference number: A97359; MUID:21608551; PMID:11743194
A/Accession: F97619
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-359 <KUR>
A/Cross-references: UNIPROT:Q8UDGO; UNIPARC:UPI00000D1DDF; GB:AE007869; PIDN:AAK87911.1;
C/Genetics:
A/Gene: AGR_C_3928
A/Map position: circular chromosome

Query Match 93.3%; Score 42; DB 2; Length 359;
Best Local Similarity 85.7%; Pred. No. 77;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPOPPPPP 7
Db 113 KPQPPPPP 119
:|||||

RESULT 11
A71416
hypothetical protein - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
A/Variety: columbia
C/Date: 03-Aug-1998 #sequence_revision 03-Aug-1998 #text_change 09-Jul-2004
C/Accession: A71416
R/Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; Dirk
P.; Nedler, H.; Wedler, E.; Wambutt, R.; Weitzenecker, T.; Pohl, T.M.; Terry, N.; Giel
avanagh, T.; Hempel, S.; Kotter, P.; Entian, K.D.; Rieger, M.; Schaeffer, M.; Funk, B.
Nature 391, 485-488, 1998
A/Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Puigdomenech
erhoff, A.; Moeres, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Rechman, S.; And
C.; Chalwatzis, N.
A/Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis thal
A/Reference number: A71400; MUID:98121113; PMID:9461215
A/Accession: A71416
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-645 <BEV>
A/Cross-references: UNIPROT:O23374; UNIPARC:UPI00000A2BC7; GB:Z97338; NID:g2244870; PID:
C/Genetics:

```

A/Map position: 4COP9-4G3B45

Query Match 93.3%; Score 42; DB 2; Length 645;
 Best Local Similarity 85.7%; Pred. No. 1.4e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPOPPPPP 7
 Db 329 KPQPPPPP 335
 :|||||

RESULT 12

IS1691
 dishevelled homolog - African clawed frog
 C/Species: Xenopus laevis (African clawed frog)
 C/Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 09-Jul-2004
 C/Accession: IS1691
 R/Sokol, S.Y.; Klingensmith, J.; Perrimon, N.; Itoh, K.
 Development 121, 1637-1647, 1995
 A/Title: Dorsalizing and neuralizing properties of Xdsh, a maternally expressed Xenopus l
 A/Reference number: IS1691; MUID:95324391; PMID:7600981
 A/Accession: IS1691
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: mRNA
 A/Residues: 1-736 <SOK>
 A/Cross-references: UNIPROT:P51142; UNIPARC:UPI00001299A9; EMBL:U31552; NID:g945109; PID:
 C/Genetics:
 A/Gene: Xdsh
 F;260-336/Domain: GLGF domain homology <GLG4>

Query Match 93.3%; Score 42; DB 2; Length 736;
 Best Local Similarity 85.7%; Pred. No. 1.5e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPOPPPPP 7
 Db 98 RPEPPPPP 104
 :|||||

RESULT 13

T30188
 Niemann-Pick C disease protein - mouse
 N/Alternate names: NPC1 protein
 C/Species: Mus musculus (house mouse)
 C/Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 09-Jul-2004
 C/Accession: T30188
 R/Loftus, S.K.; Morris, J.A.; Carstea, E.D.; Gu, J.Z.; Cummings, C.; Brown, A.; Ellison,
 Science 277, 232-235, 1997
 A/Title: Murine model of niemann-pick C disease: mutation in a cholesterol homeostasis g
 A/Reference number: Z20765; MUID:97362324; PMID:9211850
 A/Accession: T30188
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: mRNA
 A/Residues: 1-1278 <LOF>
 A/Cross-references: UNIPROT:O35604; UNIPARC:UPI00000299BC; EMBL:AF003348; NID:g2251241; I
 C/Genetics:
 A/Gene: NPC1
 A/Map position: 18

Query Match 93.3%; Score 42; DB 2; Length 1278;
 Best Local Similarity 85.7%; Pred. No. 2.6e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPOPPPPP 7
 Db 251 KPQPPPPP 257
 :|||||

RESULT 14

A46431
 tight junction-associated protein ZO-1 - mouse
 C/Species: Mus musculus (house mouse)
 C/Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004

C;Accession: A46431
R;Itoh, M.; Nagafuchi, A.; Yonemura, S.; Kitani-Yasuda, T.; Tsukita, S.; Tsukita, S.
J. Cell Biol. 121, 491-502, 1993
A;Title: The 220-KD protein colocalizing with cadherins in non-epithelial cells is identical to p120.
A;Reference number: A46431; MUID:93252986; PMID:8486731
A;Accession: A46431
A;Status: preliminary
A;Molecule type: nucleic acid
A;Residues: 1-1745 <ITO>
A;Cross-references: UNIPROT:P39447; UNIPARC:UPI00000290B4; GB:D14340; NID:g303709; PIDN:A46431
A;Experimental source: F9 cells
A;Note: sequence extracted from NCBI backbone (NCBIN:131200, NCBIP:131201)
F;27-106/Domain: GLGF domain homology <GLG1>
F;428-498/Domain: GLGF domain homology <GLG3>
F;645-794/Domain: guanylate kinase homology <GKI>

Query Match 93.3%; Score 42; DB 2; Length 1745;
Best Local Similarity 85.7%; Pred. No. 3.5e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPQPPPP 7
Db 1241 KPQPPPP 1247

RESULT 15
A83295
hypothetical protein PA2819 [imported] - Pseudomonas aeruginosa (strain PA01)
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C;Accession: A83295
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Brinkman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, J.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen.
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: A83295
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-159 <STO>
A;Cross-references: UNIPROT:Q91024; UNIPARC:UPI000000C5825; GB:AE004708; GB:AE004091; NID:A83295
A;Experimental source: strain PA01
C;Genetics:
A;Gene: PA2819

Query Match 91.1%; Score 41; DB 2; Length 159;
Best Local Similarity 85.7%; Pred. No. 47;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPQPPPP 7
Db 50 RPRPPPP 56

RESULT 16
JQ1682
infected cell protein ICP34.5 - human herpesvirus 1 (strain 17)
C;Species: human herpesvirus 1
C;Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 09-Jul-2004
C;Accession: JQ1682
R;Dolan, A.; McKie, E.; MacLean, A.R.; McGeoch, D.J.
J. Gen. Virol. 73, 971-973, 1992
A;Title: Status of the ICP34.5 gene in herpes simplex virus type 1 strain 17.
A;Reference number: JQ1682; MUID:92341080; PMID:1321882
A;Accession: JQ1682
A;Molecule type: DNA
A;Residues: 1-248 <DOL>
A;Cross-references: UNIPROT:P36313; UNIPARC:UPI000012D17E; GB:S40593
C;Genetics:
A;Gene: ICP34.5
C;Superfamily: herpesvirus infected cell protein ICP34.5

C;Keywords: tandem repeat
F;161-175/Region: 3-residue repeats (A-T-P)

Query Match 91.1%; Score 41; DB 1; Length 248;
Best Local Similarity 85.7%; Pred. No. 73;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPQPPPP 7
Db 102 RPRPPPP 108

RESULT 17
A39724
homeotic protein Hox A4 - human
N;Alternate names: homeotic protein Hox 1.4; homeotic protein Hox 1D
C;Species: Homo sapiens (man)
C;Date: 14-Feb-1992 #sequence_revision 14-Feb-1992 #text_change 09-Jul-2004
C;Accession: A39724; A60492; S15544; A32468
R;Buettnner, R.; Yim, S.O.; Hong, Y.S.; Boncinelli, E.; Tainsky, M.A.
Mol. Cell. Biol. 11, 3573-3583, 1991
A;Title: Alteration of homeobox gene expression by N-ras transformation of PA-1 human teratocarcinoma cell culture lines.
A;Reference number: A39724; MUID:91260707; PMID:1675427
A;Accession: A39724
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-320 <BUE>
A;Cross-references: UNIPROT:Q00056; UNIPARC:UPI000016AA7D; GB:M39724
R;Peverali, F.A.; D'Esposito, M.; Acampora, D.; Bunone, G.; Negri, M.; Faiella, A.; Storti, F.
Differentiation 45, 61-69, 1990
A;Title: Expression of HOX homeogenes in human neuroblastoma cell culture lines.
A;Reference number: A60492; MUID:91153613; PMID:1981366
A;Accession: A60492
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-107, 'A', 109-320 <PEV>
A;Cross-references: UNIPARC:UPI000012CF25
R;Boncinelli, E.; Acampora, D.; Pannese, M.; d'Esposito, M.; Somma, R.; Gaudino, G.; Storti, F.
Genomics 5, 250-258, 1989
A;Title: Isolation, chromosomal localization, and nucleotide sequence of the human HOX 1 gene.
A;Reference number: A32468; MUID:90007544; PMID:2571574
A;Accession: A32468
A;Molecule type: DNA
A;Residues: 215-280 <BON>
A;Cross-references: UNIPARC:UPI00001745B6
R;Ferguson-Smith, A.C.; Fienberg, A.; Ruddle, F.H.
Genomics 5, 250-258, 1989
A;Title: Organization of human class I homeobox genes.
A;Reference number: S15036; MUID:90215256; PMID:2576652
A;Accession: S15544
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 215-280 <BON>
A;Cross-references: UNIPARC:UPI00001745B7; EMBL:M28199
A;Status: This homeotic protein is expressed in embryonic nervous system.
C;Genetics:
A;Gene: GDB:HOXA4
A;Cross-references: GDB:120650; OMIM:142953
A;Map position: 7p15.3-7p15.3
C;Function:
A;Description: control of embryonic development by tissue- and stage-specific regulation
C;Superfamily: homeotic protein Hox D4; homeobox homology
C;Keywords: DNA binding; embryo; homeobox; nucleus; transcription regulation
F;216-272/Domain: homeobox homology <HOX>

Query Match 91.1%; Score 41; DB 1; Length 320;
Best Local Similarity 85.7%; Pred. No. 93;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPQPPPP 7
Db 137 QPQPPPP 143

QY 1 RPQPPPP 7
:|||||
Db 26 QPQPPPP 32

RESULT 23

T30351
mucin-like protein - Lymantria dispar nuclear polyhedrosis virus
C:Species: Lymantria dispar nuclear polyhedrosis virus, LdMNPV
C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C:Accession: T30351
R;Kuzio, J.; Pearson, M.N.; Harwood, S.H.; Funk, C.J.; Evans, J.T.; Slavicek, J.M.; Rohrer
Virology 253, 17-34, 1999
A:Title: Sequence and analysis of the genome of a baculovirus pathogenic for Lymantria d
A:Reference number: Z20836; MUID:99124785; PMID:9887315
A:Accession: T30351
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1029 <KUZ>
A:Cross-references: UNIPROT:Q9YMXO; UNIPARC:UPI000000ECOCA; EMBL:AF081810; PIDN:AAC70189

Query Match 91.1%; Score 41; DB 2; Length 1029;
Best Local Similarity 85.7%; Pred. No. 2.9e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
:|||||
Db 933 QPQPPPP 939

RESULT 24

A57278
fibrillin-2 precursor - mouse
C:Species: Mus musculus (house mouse)
C>Date: 23-Feb-1996 #sequence_revision 23-Feb-1996 #text_change 09-Jul-2004
C:Accession: A57278
R;Zhang, H.; Hu, W.; Ramirez, P.
J. Cell Biol. 129, 1165-1176, 1995
A:Title: Developmental expression of fibrillin genes suggests heterogeneity of extracell
A:Reference number: A57278; MUID:95263670; PMID:7744963
A:Accession: A57278
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-2907 <ZHA>
A:Cross-references: UNIPROT:Q61555; UNIPARC:UPI00000289B9; GB:L39790; NID:g762830; PIDN:
C:Superfamily: fibrillin; EGF homology
F;1239-1274/Domain: EGF homology <EGF1>
F;2488-2523/Domain: EGF homology <EGF>

Query Match 91.1%; Score 41; DB 2; Length 2907;
Best Local Similarity 85.7%; Pred. No. 7.7e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
:|||||
Db 29 QPQPPPP 35

RESULT 25

A54105
fibrillin-2 precursor - human
C:Species: Homo sapiens (man)
C>Date: 09-Sep-1994 #sequence_revision 09-Sep-1994 #text_change 09-Jul-2004
C:Accession: A54105; S17063; S31101
R;Zhang, H.; Apfelroth, S.D.; Hu, W.; Davis, E.C.; Sanguineti, C.; Bonadio, J.; Mecham,
J. Cell Biol. 124, 855-863, 1994
A:Title: Structure and expression of fibrillin-2, a novel microfibrillar component prefe
A:Reference number: A54105; MUID:94165150; PMID:8120105
A:Accession: A54105
A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra
A:Molecule type: mRNA
A:Residues: 1-2918 <ZHA>
A:Cross-references: UNIPROT:P35556; UNIPARC:UPI000017651B; GB:U03272

R;Lee, B.; Godfrey, M.; Vitale, E.; Hori, H.; Mattei, M.G.; Sarfarazi, M.; Tsipouras, P.
Nature 352, 330-334, 1991
A:Title: Linkage of Marfan syndrome and a phenotypically related disorder to two differ
A:Reference number: S17062; MUID:91304567; PMID:1852206
A:Accession: S17063
A:Molecule type: mRNA
A:Residues: 752-1489, 1791, 'GS', 1794-1796, 'QLI', 1922-1923, 'LD', 1926, 'P', 1928 <LEE>
A:Cross-references: UNIPARC:UPI000017651C; EMBL:X62009
R;Milewicz, D.M.
submitted to the EMBL Data Library, December 1992
A:Reference number: S31101
A:Accession: S31101
A:Molecule type: mRNA
A:Residues: 752-1407, 'R', 1409-1489, 1791, 'GS', 1794-1796, 'QLI', 1922-1923, 'LD', 1926, 'P', 1928
A:Cross-references: UNIPARC:UPI000000069B; EMBL:X62009
C:Genetics:
A:Gene: GDB:FBN2
A:Cross-references: GDB:128122; OMIM:121050
A:Map position: 5q23-5q31
C:Superfamily: fibrillin; EGF homology
C:Keywords: extracellular protein
F;1-29/Domain: signal sequence #status predicted <SIG>
F;30-2918/Product: fibrillin-2 #status predicted <MAT>
F;1245-1280/Domain: EGF homology <EGF1>
F;1970-2013/Domain: EGF homology <EGF>

Query Match 91.1%; Score 41; DB 2; Length 2918;
Best Local Similarity 85.7%; Pred. No. 7.8e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
:|||||
Db 29 QPQPPPP 35

RESULT 26

A46068
Huntington disease-associated protein - human
C:Species: Homo sapiens (man)
C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
C:Accession: A46068; I54337
R;MacDonald, M.E.; Ambrose, C.M.; Duyao, M.P.; Myers, R.H.; Lin, C.; Srinidhi, L.; Barnes
J.F.; Bates, G.P.; Baxendale, S.; Hummerich, H.; Kirby, S.; North, M.; Youngman, S.; Mott
O'Donovan, M.C.; Riba-Ramirez, L.; Shah, M.; Stanton, V.P.; Strobel, S.A.; Draths, K.M.
Cell 72, 971-983, 1993
A:Authors: Wales, J.L.; Dervan, P.; Houseman, D.E.; Altherr, M.; Shiang, R.; Thompson, L.,
d, K.; Collins, F.S.; Snell, R.; Holloway, T.; Gillespie, K.; Datson, N.; Shaw, D.; Harpe
A:Title: A novel gene containing a trinucleotide repeat that is expanded and unstable on
A:Reference number: A46068; MUID:93208892; PMID:8458085
A:Accession: A46068
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-3144 <MAC>
A:Cross-references: UNIPROT:P42858; UNIPARC:UPI0000146F05; GB:L12392
R;Lin, B.; Rommens, J.M.; Graham, R.K.; Kalchman, M.; MacDonald, H.; Nasir, J.; Delaney,
Hum. Mol. Genet. 2, 1541-1545, 1993
A:Title: Differential 3' polyadenylation of the Huntington disease gene results in two m
A:Reference number: I54337; MUID:94093536; PMID:7903579
A:Accession: I54337
A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA
A:Residues: 2563-3144 <RES>
A:Cross-references: UNIPARC:UPI000016A8BF; GB:L20431; NID:g398028; PIDN:AAA52702.1; PID:
C:Genetics:
A:Gene: GDB:HD
A:Cross-references: GDB:119307; OMIM:143100
A:Map position: 4p16.3-4p16.3

Query Match 91.1%; Score 41; DB 2; Length 3144;
Best Local Similarity 85.7%; Pred. No. 8.4e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7

```

Db          66 QPQPPPP 72
:|||||
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

RESULT 27
A96650
protein F2401.6 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C;Accession: A96650
R;Theologias, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, B.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: A96650
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-70 <STO>
A;Cross-references: UNIPROT:Q9MAV4; UNIPARC:UPI00000A2B99; GB:AE005173; NID:G7940276; PIDN:
C;Genetics:
A;Map Position: 1
A;Gene: F2401.6
C;Superfamily: hydroxyproline-rich glycoprotein

Query Match 88.9%; Score 40; DB 2; Length 70;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 PQPPPP 7
|||||
Db 35 PQPPPP 40

RESULT 28
B48232
cysteine-rich extensin-like protein 2 precursor - common tobacco
C;Species: Nicotiana tabacum (common tobacco)
C;Date: 26-May-1994 #sequence_revision 26-May-1994 #text_change 09-Jul-2004
C;Accession: B48232; PQ0474; S24616
R;Wu, H.; Zou, J.; May, B.; Gu, Q.; Cheung, A.Y.
Proc. Natl. Acad. Sci. U.S.A. 90, 6829-6833, 1993
A;Title: A tobacco gene family for flower cell wall proteins with a proline-rich domain
A;Reference number: A48232; MUID:93342083; PMID:8341705
A;Accession: B48232
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-196 <WUA>
A;Cross-references: UNIPROT:Q08195; UNIPARC:UPI00000A3E44; GB:LI13440; NID:G310924; PIDN:
R;de S Goldman, M.H.; Pezzotti, M.; Seurinck, J.; Mariani, C.
Plant Cell 4, 1041-1051, 1992
A;Title: Developmental expression of tobacco pistil-specific genes encoding novel extens
A;Reference number: PQ0474; MUID:93005740; PMID:1392607
A;Accession: B48232
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 'MAG', 1-105 <GOL>
A;Cross-references: UNIPARC:UPI0000177E65; EMBL:Z14014
A;Experimental source: stigma, style; strain Petit Havana SRI
C;Genetics:
A;Gene: CELP-2
C;Superfamily: glutelin
C;Keywords: cell wall; extracellular matrix; fertilization
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-196/Product: cysteine-rich extensin-like protein 2 #status experimental <MAT>

* Query Match 88.9%; Score 40; DB 2; Length 196;
Best Local Similarity 85.7%; Pred. No. 78;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RQPPPPP 7
|||||
Db 100 RPSPPPP 106

RESULT 29
A48232
cysteine-rich extensin-like protein 1 precursor - common tobacco
C;Species: Nicotiana tabacum (common tobacco)
C;Date: 26-May-1994 #sequence_revision 26-May-1994 #text_change 09-Jul-2004
C;Accession: A48232; PQ0475; S24617
R;Wu, H.; Zou, J.; May, B.; Gu, Q.; Cheung, A.Y.
Proc. Natl. Acad. Sci. U.S.A. 90, 6829-6833, 1993
A;Title: A tobacco gene family for flower cell wall proteins with a proline-rich domain
A;Reference number: A48232; MUID:93342083; PMID:8341705
A;Accession: A48232
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-209 <WUA>
A;Cross-references: UNIPROT:Q08194; UNIPARC:UPI00000ABB3F; GB:LI13439; NID:G310922; PIDN:
R;de S Goldman, M.H.; Pezzotti, M.; Seurinck, J.; Mariani, C.
Plant Cell 4, 1041-1051, 1992
A;Title: Developmental expression of tobacco pistil-specific genes encoding novel extensin
A;Reference number: PQ0474; MUID:93005740; PMID:1392607
A;Accession: PQ0475
A;Molecule type: mRNA
A;Residues: 39-209 <GOL>
A;Cross-references: UNIPARC:UPI000016DB8A; EMBL:Z14020; NID:G19918; PID:G19919
A;Experimental source: stigma, style; strain Petit Havana SRI
C;Superfamily: glutelin
C;Keywords: cell wall; extracellular matrix; fertilization; glycoprotein
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-209/Product: cysteine-rich extensin-like protein 1 #status experimental <MAT>
F;146/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 88.9%; Score 40; DB 2; Length 209;
Best Local Similarity 85.7%; Pred. No. 83;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RQPPPPP 7
|||||
Db 115 RPSPPPP 121

RESULT 30
S34163
homeotic protein Hoxd-3 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
C;Accession: S34163
R;Bedford, M.; Petr, L.
submitted to the EMBL Data Library, June 1993
A;Reference number: S34163
A;Accession: S34163
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-215 <BED>
A;Cross-references: UNIPROT:Q61685; UNIPARC:UPI00000028C7F; EMBL:X73572; NID:G313248; PIDN:
C;Genetics:
A;Introns: 165/3

Query Match 88.9%; Score 40; DB 2; Length 215;
Best Local Similarity 100.0%; Pred. No. 85;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 PQPPPP 7
|||||
Db 101 PQPPPP 106

RESULT 31
```

T07176
extensin homolog - potato (fragment)
C;Species: Solanum tuberosum (potato)
C;Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 09-Jul-2004
C;Accession: T07176
R;MacLeod, M.R.
submitted to the EMBL Data Library, December 1997
A;Reference number: 215977
A;Accession: T07176
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-221 <MAC>
A;Cross-references: UNIPROT:Q49946; UNIPARC:UPI00000ACA27; EMBL:AJ003220; NID:e1251331;
A;Experimental source: cv. Record; swelling stolon
C;Superfamily: glutelin

Query Match 88.9%; Score 40; DB 2; Length 221;
Best Local Similarity 85.7%; Pred. No. 87;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RPQPPPP 7
Db 50 RPSPPPP 56

RESULT 32
S54156
extensin-like protein - cowpea (fragment)
C;Species: Vigna unguiculata (cowpea)
C;Date: 08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change 09-Jul-2004
C;Accession: S54156
R;Arsenijevic-Maksimovic, I.; Broughton, W.J.; Krause, A.
submitted to the EMBL Data Library, April 1995
A;Description: A class of root-hair specific extensins involved in rhizobium/legume inte
A;Reference number: S54155
A;Accession: S54156
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-242 <ARS>
A;Cross-references: UNIPROT:Q43687; UNIPARC:UPI000009E1E7; EMBL:X86029; NID:g791147; PID
C;Superfamily: hydroxyproline-rich glycoprotein

Query Match 88.9%; Score 40; DB 2; Length 242;
Best Local Similarity 85.7%; Pred. No. 95;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RPQPPPP 7
Db 128 RPSPPPP 134

RESULT 33
C86480
33.2K hypothetical protein - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C;Accession: C86480
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: C86480
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-302 <STO>
A;Cross-references: UNIPROT:Q9C8B9; UNIPARC:UPI00000A8A47; GB:AE005172; NID:g11055830; H

C;Genetics:
A;Map position: 1

Query Match 88.9%; Score 40; DB 2; Length 302;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 PQPPPP 7
Db 82 PQPPPP 87

RESULT 34
JQ1386
hypothetical 33K protein - pea enation mosaic virus
C;Species: pea enation mosaic virus, PEMV
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004
C;Accession: JQ1386
R;Demler, S.A.; de Zoeten, G.A.
J. Gen. Virol. 72, 1819-1834, 1991
A;Title: The nucleotide sequence and luteovirus-like nature of RNA 1 of an aphid non-tra
A;Reference number: JQ1382; MUID:91341468; PMID:1875194
A;Accession: JQ1386
A;Molecule type: genomic RNA
A;Residues: 1-303 <DEM>
A;Cross-references: UNIPROT:Q84711; UNIPARC:UPI00000F0B4D; GB:L04573; NID:g294105; PIDN:
A;Experimental source: strain WSG
A;Note: 65-Asp was also found
C;Genetics:
A;Map position: segment RNA1

Query Match 88.9%; Score 40; DB 2; Length 303;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 PQPPPP 7
Db 14 PQPPPP 19

RESULT 35
B29350
env polyprotein, retrovirus-related - mouse (fragment)
N;Contains: coat protein gp70
C;Species: Mus musculus (house mouse)
C;Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 15-Nov-1996
C;Accession: B29350
R;Levy, D.E.; Lerner, R.A.; Wilson, M.C.
J. Virol. 56, 691-700, 1985
A;Title: Normal expression of polymorphic endogenous retroviral RNA containing segments
A;Reference number: A29350; MUID:86062897; PMID:2415714
A;Accession: B29350
A;Molecule type: mRNA
A;Residues: 1-312 <LEV>
A;Cross-references: UNIPARC:UPI000017863E
A;Experimental source: clone E1
A;Note: the authors translated the codon CTA for residue 147 as Lys
C;Superfamily: type C retrovirus env polyprotein
C;Keywords: coat protein; glycoprotein; polyprotein; spike protein; transmembrane protei
F;22-312/Product: coat protein gp70 (fragment) #status predicted <GFS>
F;43,58,297/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 88.9%; Score 40; DB 2; Length 312;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 PQPPPP 7
Db 258 PQPPPP 263

RESULT 36

E70519
 probable oxidoreductase - Mycobacterium tuberculosis (strain H37RV)
 C;Species: Mycobacterium tuberculosis
 C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
 C;Accession: E70519
 R;Cole, S.T.; Broech, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, K.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998
 A;Authors: Squares, R.; Sultson, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
 A;Reference number: A70500; MUID:98295987; PMID:9634230
 A;Accession: E70519
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-334 <COL>
 A;Cross-references: UNIPROT:O07721; UNIPARC:UPI000004B72C; GB:Z97193; GB:AL123456; NID:98295987
 A;Experimental source: strain H37RV
 C;Genetics:
 A;Gene: fadB5
 C;Superfamily: alcohol dehydrogenase; long-chain alcohol dehydrogenase homology
 F;26-323/Domain: long-chain alcohol dehydrogenase homology <LADH>
 Query Match 88.9%; Score 40; DB 2; Length 334;
 Best Local Similarity 85.7%; Pred. No. 1.3e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 RPQPPPP 7
 Db 20 RPDPPPP 26
 RESULT 37
 T33836
 hypothetical protein F54A5.3a - Caenorhabditis elegans
 C;Species: Caenorhabditis elegans
 C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
 C;Accession: T33836
 R;Jones, K.; Graves, T.; Ozerky, P.; Wilson, R.
 submitted to the EMBL Data Library, November 1998
 A;Description: The sequence of C. elegans cosmid F54A5.
 A;Reference number: Z21420
 A;Accession: T33836
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-334 <ION>
 A;Cross-references: UNIPROT:Q9TYT3; UNIPARC:UPI0000017BA0D; EMBL:AF106584; PIDN:AAC78222.
 A;Experimental source: strain Bristol N2; clone F54A5
 C;Genetics:
 A;Gene: CBSP:F54A5.3a
 A;Map position: 1
 A;Introns: 22/3; 46/2; 70/3; 234/2
 Query Match 88.9%; Score 40; DB 2; Length 334;
 Best Local Similarity 85.7%; Pred. No. 1.3e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 RPQPPPP 7
 Db 266 RPSPPPP 272
 RESULT 38
 JC5833
 transcription factor E2F-5 - human
 N;Alternate names: p130/p107-binding protein E2F-5
 C;Species: Homo sapiens (man)
 C;Date: 20-Feb-1998 #sequence_revision 13-Mar-1998 #text_change 05-Oct-2004
 C;Accession: JC5833; I38878
 R;Vaishnav, Y.N.; Vaishnav, M.Y.; Pant, V.
 Biochem. Biophys. Res. Commun. 242, 586-592, 1998
 A;Title: The molecular and functional characterization of E2F-5 transcription factor.
 A;Reference number: JC5833; MUID:98125536; PMID:9464260

A;Accession: JC5833
 A;Status: nucleic acid sequence not shown
 A;Molecule type: mRNA
 A;Residues: 1-346 <VAI>
 A;Cross-references: UNIPROT:Q15329; UNIPARC:UPI000002E057; EMBL:Z78409; NID:gl592737
 R;Sardet, C.; Vidal, M.; Cobrinik, D.; Geng, Y.; Onufryk, C.; Chen, A.; Weinberg, R.A. Proc. Natl. Acad. Sci. U.S.A. 92, 2403-2407, 1995
 A;Title: E2F-4 and E2F-5, two members of the E2F family, are expressed in the early phase of the cell cycle
 A;Reference number: I38877; MUID:95199358; PMID:7892279
 A;Accession: I38878
 A;Status: preliminary
 A;Molecule type: mRNA
 A;Residues: 1-294,296-346 <SAR>
 A;Cross-references: UNIPARC:UPI000016A135; EMBL:U15642; NID:g758415; PIDN:AAC50120.1; PII:15642
 C;Genetics:
 A;Gene: GDB:E2F5
 A;Cross-references: GDB:578742; OMIM:600967
 A;Map position: 16q22.1-16q22.1
 C;Superfamily: transcription factor E2F5
 C;Keywords: cell cycle control; DNA binding
 F;49-116/Domain: DNA-binding #status predicted <DNA>
 Query Match 88.9%; Score 40; DB 2; Length 346;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 PQPPPP 7
 Db 33 PQPPPP 38
 RESULT 39
 T49791
 hypothetical protein B9J10.280 [imported] - Neurospora crassa
 C;Species: Neurospora crassa
 C;Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 09-Jul-2004
 C;Accession: T49791
 R;Schulte, U.; Aign, V.; Hobeisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura, submitted to the Protein Sequence Database, May 2000
 A;Reference number: Z25022
 A;Accession: T49791
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-349 <SCH>
 A;Cross-references: UNIPROT:Q9P555; UNIPARC:UPI000006B94E; EMBL:AL356324; GSPDB:GN00116;
 A;Experimental source: BAC clone B9J10; strain OR74A
 C;Genetics:
 A;Gene: NCSP:B9J10.280
 A;Map position: 6
 Query Match 88.9%; Score 40; DB 2; Length 349;
 Best Local Similarity 85.7%; Pred. No. 1.4e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 RPQPPPP 7
 Db 181 RPHPPPP 187
 RESULT 40
 VCVWB6
 env polyprotein precursor - Friend spleen focus-forming virus (strain BB6)
 N;Alternate names: coat polyprotein
 N;Contains: knob protein gp70; spike protein p15E
 C;Species: Friend spleen focus-forming virus
 C;Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 09-Jul-2004
 C;Accession: A41995
 R;Majumdar, M.K.; Cho, C.L.; Fox, M.T.; Eckner, K.L.; Kozak, S.; Kabat, D.; Geib, R.W. J. Virol. 66, 3652-3660, 1992
 A;Title: Mutations in the env gene of Friend spleen focus-forming virus overcome Fv-2(r)-mediated restriction
 A;Reference number: A41995; MUID:92260637; PMID:1583724
 A;Accession: A41995
 A;Molecule type: DNA

A;Residues: 1-356 <MAJ>
A;Cross-references: UNIPROT:P31793; UNIPARC:UPI000012A007; GB:M90673
C;Genetics:

A;Gene: env
C;Superfamily: type C retrovirus env polyprotein
C;Keywords: coat protein; glycoprotein; polyprotein
F;1-33/Domain: signal sequence #status predicted <SIG>
F;34-356/Product: env polyprotein #status predicted <ENV>
F;34-330/Product: knob protein gp70 #status predicted <KPG>
F;331-356/Product: spike protein p15E #status predicted <SPP>
F;335-351/Region: hydrophobic #status predicted
F;43,58,296,328/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 88.9%; Score 40; DB 1; Length 356;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 PQPPPP 7
Db 257 PQPPPP 262

RESULT 41

S31428
envelope protein - spleen focus-forming virus
C;Species: spleen focus-forming virus, SFFV
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
C;Accession: S31428
R;Cho, C.L.; Majumdar, M.K.; Gaib, R.W.
submitted to the EMBL Data Library, May 1992
A;Reference number: S31428

A;Accession: S31428
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-356 <CHO>
A;Cross-references: UNIPROT:Q87044; UNIPARC:UPI000010210C; EMBL:X66721; NID:g61959; PIDN:
C;Superfamily: type C retrovirus env polyprotein

Query Match 88.9%; Score 40; DB 2; Length 356;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 PQPPPP 7
Db 257 PQPPPP 262

RESULT 42

VCMVSR
env polyprotein - Rauscher spleen focus-forming virus
N;Alternate names: coat polyprotein
N;Contains: coat protein gp70; coat protein p15E
C;Species: Rauscher spleen focus-forming virus
C;Date: 28-Aug-1985 #sequence_revision 28-Aug-1985 #text_change 09-Jul-2004
C;Accession: A03988
R;Bestwick, R.K.; Boswell, B.A.; Kabat, D.
J. Virol. 51, 695-705, 1984

A;Title: Molecular cloning of biologically active Rauscher spleen focus-forming virus an
A;Reference number: A93001; MUID:84292446; PMID:6088793

A;Accession: A03988
A;Molecule type: DNA
A;Residues: 1-408 <BES>

A;Cross-references: UNIPROT:P03389; UNIPARC:UPI000012A048; GB:K02375; NID:g331998; PIDN:
C;Genetics:

A;Gene: env
C;Superfamily: type C retrovirus env polyprotein
C;Keywords: coat protein; glycoprotein; polyprotein
F;1-33/Domain: leader peptide #status predicted <LDP>
F;34-333/Product: coat protein gp70 #status predicted <CPA>
F;334-408/Product: coat protein p15E #status predicted <CPE>
F;43,58,297,329,378/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 88.9%; Score 40; DB 1; Length 408;

Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 PQPPPP 7
Db 258 PQPPPP 263

RESULT 43

S50899
betaB inhibin precursor - bovine
C;Species: Bos primigenius taurus (cattle)
C;Date: 01-Aug-1995 #sequence_revision 01-Sep-1995 #text_change 09-Jul-2004
C;Accession: S50899
R;Thompson, D.A.; Cronin, C.N.; Martin, F.
Eur. J. Biochem. 226, 751-764, 1994
A;Title: Genomic cloning and sequence analyses of the bovine alpha-, beta(A)- and beta(B)-
Y DNase I footprinting.

A;Reference number: S50897; MUID:95112839; PMID:7813465

A;Accession: S50899
A;Status: preliminary
A;Molecule type: DNA

A;Residues: 1-408 <THO>
A;Cross-references: UNIPROT:P42917; UNIPARC:UPI000012D425; EMBL:U16240

C;Genetics:

A;Introns: 151/1

C;Superfamily: inhibin

Query Match 88.9%; Score 40; DB 2; Length 408;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 PQPPPP 7
Db 39 PQPPPP 44

RESULT 44

A21170
gp55 protein - Friend spleen focus-forming virus
C;Species: Friend spleen focus-forming virus
C;Date: 14-Dec-1990 #sequence_revision 14-Dec-1990 #text_change 08-Apr-1994
C;Accession: A21170
R;Amanuma, H.; Katori, A.; Obata, M.; Sagata, N.; Ikawa, Y.
Proc. Natl. Acad. Sci. U.S.A. 80, 3913-3917, 1983

A;Title: Complete nucleotide sequence of the gene for the specific glycoprotein (gp55) o
A;Reference number: A21170; MUID:83247388; PMID:6306650

A;Accession: A21170
A;Status: preliminary
A;Molecule type: DNA

A;Residues: 1-409 <AMA>
A;Cross-references: UNIPARC:UPI0000178640

C;Superfamily: type C retrovirus env polyprotein

Query Match 88.9%; Score 40; DB 2; Length 409;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 PQPPPP 7
Db 257 PQPPPP 262

RESULT 45

S34164
homeotic protein Hoxd-3 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 05-Oct-2004
C;Accession: S34164
R;Bedford, M.; Petr, L.
submitted to the EMBL Data Library, June 1993
A;Reference number: S34163
A;Accession: S34164

```
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-411 <BRD>
A;Cross-references: UNIPROT:P09027; UNIPARC:UPI0000016CE02; EMBL:X73573; NID:g313252; PID:
C;Species: Caenorhabditis elegans
C;Keywords: DNA binding; homeobox; nucleus; transcription regulation
F;180-236/Domain: homeobox homology <HOX>

Query Match      88.9%; Score 40; DB 2; Length 411;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 PQPPPP 7
DB      101 PQPPPP 106

RESULT 46
T04520
Hypothetical protein F16A16.100 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 09-Jul-2004
C;Accession: T04520
R;Bevan, M.; Brandt, P.; Dese, S.; Jarke, D.; Scharfe, M.; Schon, O.; Hoheisel, J.; Newe
submitted to the Protein Sequence Database, February 1999
A;Reference number: Z15376
A;Accession: T04520
A;Molecule type: DNA
A;Residues: 1-413 <BRV>
A;Cross-references: UNIPROT:Q99VU6; UNIPARC:UPI000000A6306; EMBL:AL035353
A;Experimental source: cultivar Columbia; BAC clone F16A16
C;Genetics:
A;Map position: 4
A;Introns: 22/2; 262/3; 287/3; 309/3; 331/3
A;Note: F16A16.100

Query Match      88.9%; Score 40; DB 2; Length 413;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 PQPPPP 7
DB      69 PQPPPP 74

RESULT 47
S27198
homeotic protein Hox D3 - human
N;Alternate names: homeotic protein Hox 4A
C;Species: Homo sapiens (man)
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
C;Accession: S27198
R;Taniguchi, Y.; Fujii, A.; Moriuchi, T.
Biochim. Biophys. Acta 1132, 332-334, 1992
A;Title: Cloning and sequencing of the human homeobox gene HOX4A.
A;Reference number: S27198; MUID:93041940; PMID:1358204
A;Accession: S27198
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-416 <TAN>
A;Cross-references: UNIPROT:P31249; UNIPARC:UPI000016AA89; EMBL:D11117; NID:g219879; PID:
C;Genetics:
A;Introns: 165/1
A;Superfamily: homeotic protein Hox B3; homeobox homology
C;Keywords: DNA binding; homeobox; nucleus; transcription regulation
F;179-235/Domain: homeobox homology <HOX>

Query Match      88.9%; Score 40; DB 2; Length 416;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 PQPPPP 7
DB      101 PQPPPP 106
```

RESULT 48

T34279

hypothetical protein F46H5.6 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004

C;Accession: T34279

R;Nhan, M.

submitted to the EMBL Data Library, November 1995

A;Description: The sequence of C. elegans cosmid F46H5.

A;Reference number: Z21498

A;Accession: T34279

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-416 <NHA>

A;Cross-references: UNIPROT:P90878; UNIPARC:UPI000007BC53; EMBL:U41543; PIDN:AAB37027.1;

A;Experimental source: strain Bristol N2; clone F46H5

C;Genetics:

A;Gene: CBSP:F46H5.6

A;Map position: X

A;Introns: 231/3; 260/3; 325/3; 352/2; 385/2

Query Match

88.9%; Score 40; DB 2; Length 416;

Best Local Similarity 100.0%; Pred. No. 1.6e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 PQPPPP 7

DB 39 PQPPPP 44

RESULT 49

S47539

homeotic protein Hox D3 - mouse

N;Alternate names: homeobox-containing protein Hox-4.1

C;Species: Mus musculus (house mouse)

C;Date: 27-Jan-1995 #sequence_revision 27-Jan-1995 #text_change 09-Jul-2004

C;Accession: S47539; JC6144

R;Brown, W.M.; Zhou, L.; Taylor, G.R.

Biochim. Biophys. Acta 1219, 219-222, 1994

A;Title: The nucleotide sequence of the murine Hox-D3 (Hox-4.1) gene reveals extensive i

A;Reference number: S47539; MUID:94368862; PMID:7916214

A;Accession: S47539

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-417 <BRO>

A;Cross-references: UNIPROT:P09027; UNIPARC:UPI0000026946; EMBL:U03496

R;Tan, D.P.; Shao, X.; Pu, L.; Guo, V.; Nirenberg, M.

Proc. Natl. Acad. Sci. U.S.A. 93, 8247-8252, 1996

A;Title: Sequence and expression of the murine Hoxd-3 homeobox gene.

A;Reference number: JC6144; MUID:96323206; PMID:8710855

A;Contents: embryo

A;Accession: JC6144

A;Molecule type: mRNA

A;Residues: 1-417 <TAN>

A;Cross-references: UNIPARC:UPI0000026946; GB:U56400; NID:g1398928; PIDN:AAC52779.1; PID:

C;Comment: This protein is a homeodomain protein with SH3 domain, and it controls the rat

C;Genetics:

A;Introns: 166/1

A;Gene: Hoxd-3

C;Superfamily: homeotic protein Hox B3; homeobox homology

C;Keywords: DNA binding; homeobox; nucleus; transcription regulation

F;180-236/Domain: homeobox homology <HOX>

Query Match

88.9%; Score 40; DB 2; Length 417;

Best Local Similarity 100.0%; Pred. No. 1.6e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 PQPPPP 7

DB 101 PQPPPP 106

RESULT 50
A83184
Probable protein methyltransferase PA3706 [imported] - Pseudomonas aeruginosa (strain PA
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C;Accession: A83184
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Bu
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A;Reference number: A82950; MUID:2043737; PMID:10984043
A;Accession: A83184
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-422 <STO>
A;Cross-references: UNIPROT:Q9HXT5; UNIPARC:UPI000000CSACD; GB:AE004789; GB:AE004091; NID
A;Experimental source: strain PA01
C;Genetics:
A;Gene: PA3706

Query Match 88.9%; Score 40; DB 2; Length 422;
Best Local Similarity 85.7%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RPQPPPP 7
Db 292 RPSPPPP 298

Search completed: April 6, 2006, 09:40:09
Job time : 17.7105 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 6, 2006, 09:23:24 ; Search time 85,1053 Seconds
(without alignments)
58.030 Million cell updates/sec

Title: US-10-632-388-305

Perfect score: 45

Sequence: 1 RQPPPP 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Database : Uniprot_05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	45	100.0	66	2	Q61IP9_DROME
2	45	100.0	73	2	Q5VP16_ORYSA
3	45	100.0	119	2	Q5VPA7_ORYSA
4	45	100.0	125	2	Q6LDK2_9MURI
5	45	100.0	132	2	Q6VZB3_ORYSA
6	45	100.0	165	2	Q7PBE9_ORYSA
7	45	100.0	208	2	Q654U7_ORYSA
8	45	100.0	270	2	Q9ZA63_NEIME
9	45	100.0	291	2	Q627W0_ORYSA
10	45	100.0	300	2	Q8LIW2_ORYSA
11	45	100.0	317	2	Q6MP85_BDEBA
12	45	100.0	328	2	Q4T9Z6_TETNG
13	45	100.0	398	2	Q9LHG8_ARATH
14	45	100.0	416	2	Q9C2D6_NEUCR
15	45	100.0	420	2	Q9VZC2_DROME
16	45	100.0	449	2	Q9PZ18_GVXN
17	45	100.0	551	2	Q4P0X1_USTMA
18	45	100.0	587	2	Q525T5_MAGGR
19	45	100.0	624	2	Q4T445_TETNG
20	45	100.0	669	2	Q5H9B0_HUMAN
21	45	100.0	687	2	Q4G0U1_HUMAN
22	45	100.0	704	1	SYN1_RAT
23	45	100.0	704	2	Q95XQ7_CABEL
24	45	100.0	705	1	SYN1_HUMAN
25	45	100.0	706	1	SYN1_BOVIN
26	45	100.0	706	1	SYN1_MOUSE
27	45	100.0	721	2	Q51L36_MAGGR
28	45	100.0	1412	2	Q4QF52_LEIMA
29	45	100.0	1480	2	Q9G0Q4_HUMAN
30	45	100.0	1798	1	MAST2_HUMAN
31	45	100.0	2301	2	Q9ATK5_CHLRE

32	45	100.0	5262	1	MLL2_HUMAN
33	42	93.3	107	2	Q67YU2_ARATH
34	42	93.3	139	2	Q9SVS5_ARATH
35	42	93.3	169	2	Q5N912_ORYSA
36	42	93.3	246	2	Q620S6_ORYSA
37	42	93.3	279	2	Q6ZQX5_HUMAN
38	42	93.3	287	2	Q96KT3_HUMAN
39	42	93.3	309	2	Q7XPK1_ORYSA
40	42	93.3	320	2	Q8UDG0_AGR5
41	42	93.3	355	2	Q4IEH9_GIBZE
42	42	93.3	359	2	Q7CXQ3_AGR5
43	42	93.3	377	2	Q4S7B1_TETNG
44	42	93.3	464	2	Q61C10_CABBR
45	42	93.3	493	2	Q9IUV1_PNECA
46	42	93.3	496	2	Q9RXK6_STRCO
47	42	93.3	542	2	Q615B7_ORYSA
48	42	93.3	549	2	Q9FMF8_ARATH
49	42	93.3	579	2	Q4RP94_TETNG
50	42	93.3	619	2	Q6AHW2_PNECA
51	42	93.3	641	1	XK36_HUMAN
52	42	93.3	645	2	Q23374_ARATH
53	42	93.3	670	2	Q8RVE5_ORYSA
54	42	93.3	733	2	Q5EAW6_XENLA
55	42	93.3	736	1	DVL2_XENLA
56	42	93.3	757	2	Q91YW2_MOUSE
57	42	93.3	760	2	Q86A14_DICDI
58	42	93.3	782	1	SOX30_MOUSE
59	42	93.3	782	2	Q5SSN4_MOUSE
60	42	93.3	858	2	Q415L0_GIBZE
61	42	93.3	873	2	Q41BP4_GIBZE
62	42	93.3	902	2	Q6AHS3_PNECA
63	42	93.3	902	2	Q6AHV9_PNECA
64	42	93.3	933	2	Q9BL41_CABEL
65	42	93.3	947	2	Q6AHS6_PNECA
66	42	93.3	1158	2	Q4QGW6_LEIMA
67	42	93.3	1264	2	Q6ZU65_HUMAN
68	42	93.3	1276	2	Q8MI49_FELCA
69	42	93.3	1276	2	Q8MKD8_FELCA
70	42	93.3	1276	2	Q9NOQ0_FELCA
71	42	93.3	1277	1	NPC1_PIG
72	42	93.3	1277	2	Q9GLC9_BOVIN
73	42	93.3	1277	2	Q7TMD4_MOUSE
74	42	93.3	1278	1	NPC1_HUMAN
75	42	93.3	1278	1	NPC1_MOUSE
76	42	93.3	1286	2	Q9IT75_RABIT
77	42	93.3	1289	2	Q59GR1_HUMAN
78	42	93.3	1313	2	Q80WC1_MOUSE
79	42	93.3	1393	2	Q81PP9_DROME
80	42	93.3	1417	2	Q7YU09_DROME
81	42	93.3	1461	2	Q4QC44_LEIMA
82	42	93.3	1621	2	Q9V1K7_DROME
83	42	93.3	1745	1	ZOI_MOUSE
84	42	93.3	2691	2	Q61B74_CABBR
85	42	93.3	2761	2	Q4P226_USTMA
86	42	93.3	3467	2	Q81218_PLAF7
87	42	93.3	3903	1	MLL3_MOUSE
88	41	91.1	33	2	Q02832_9PRIM
89	41	91.1	58	2	Q68GZ9_BOVIN
90	41	91.1	58	2	Q68GY0_CANLA
91	41	91.1	70	2	Q62312_ORYSA
92	41	91.1	77	2	Q5QMJ7_ORYSA
93	41	91.1	95	1	BCD_DROSU
94	41	91.1	101	2	Q6WZC9_BOVIN
95	41	91.1	104	2	Q5DM45_BOVIN
96	41	91.1	107	2	Q03900_9VIRU
97	41	91.1	127	2	Q69JN5_ORYSA
98	41	91.1	131	1	ASIP_CANFA
99	41	91.1	131	1	ASIP_VULVU
100	41	91.1	133	2	Q5QLQ9_ORYSA
101	41	91.1	135	1	ASIP_FELCA
102	41	91.1	135	2	Q624C7_ORYSA
103	41	91.1	150	2	Q51UH7_MAGGR
104	41	91.1	150	2	Q8C005_MOUSE

014686	homo sapien
067YU2	arabidopsis
Q9SVS5	arabidopsis
Q5N912	oryza sativ
Q620S6	oryza sativ
Q6ZQX5	homo sapien
Q96KT3	homo sapien
Q7XPK1	oryza sativ
Q8UDG0	agrobacteri
Q4IEH9	gibberella
Q7CXQ3	agrobacteri
Q4S7B1	tetradodon n
Q61C10	caenorhabdi
Q9IUV1	pneumocysti
Q9RXK6	streptomyce
Q615B7	oryza sativ
Q9FMF8	arabidopsis
Q4RP94	tetradodon n
Q6AHW2	pneumocysti
Q5GH73	homo sapien
Q23374	arabidopsis
Q8RVE5	oryza sativ
Q5EAW6	xenopus lae
P51142	xenopus lae
Q91YW2	mus musculu
Q86A14	dictyosteli
Q8CGW4	mus musculu
Q5SSN4	mus musculu
Q415L0	gibberella
Q41BP4	gibberella
Q6AHS3	pneumocysti
Q6AHV9	pneumocysti
Q9BL41	caenorhabdi
Q6AHS6	pneumocysti
Q4QGW6	leishmania
Q6ZU65	homo sapien
Q8MI49	felis silve
Q8MKD8	felis silve
Q9NOQ0	felis silve
P56941	sus scrofa
Q9GLC9	bos taurus
Q7TMD4	mus musculu
O15118	homo sapien
Q35604	mus musculu
Q9IT75	oryctolagus
Q59GR1	homo sapien
Q80WC1	mus musculu
Q81PP9	drosophila
Q7YU09	drosophila
Q4QC44	leishmania
Q9V1K7	drosophila
P39447	mus musculu
Q618T4	caenorhabdi
Q68GY0	canis lupus
Q4P226	ustilago ma
Q81218	plasmodium
Q8BTH4	mus musculu
O02832	gorilla gor
Q68GZ9	canis lupus
Q68GY0	canis lupus
Q62312	oryza sativ
Q5QMJ7	oryza sativ
P55924	drosophila
Q6WZC9	bos taurus
Q5DM45	bos taurus
O03900	bacterioph
Q69JN5	oryza sativ
Q5UK76	canis famil
P79407	vulpes vulp
Q5QLQ9	oryza sativ
Q865F0	felis silve
Q624C7	oryza sativ
Q51UH7	magnaporthe
Q8C005	mus musculu

105	41	91.1	159	2	Q91024_PSEAB	Q91024_pseudomonas	178	41	91.1	822	2	Q7XCT3_ORYSA	Q7xct3 oryza sativ
106	41	91.1	176	2	Q8LM61_ORYSA	Q8lm61 oryza sativ	179	41	91.1	822	2	Q94HZ3_ORYSA	Q94hz3 oryza sativ
107	41	91.1	176	2	Q5ZB80_ORYSA	Q5zb80 oryza sativ	180	41	91.1	834	2	Q5SPX0_CRYNE	Q5spx0 cryptococcus
108	41	91.1	181	2	Q7QYV0_GIALA	Q7qyv0 giardia lam	181	41	91.1	834	2	Q5KDJ0_CRYNE	Q5kdj0 cryptococcus
109	41	91.1	187	2	Q654G3_ORYSA	Q654g3 oryza sativ	182	41	91.1	837	1	FA40A_HUMAN	Q5vsl9 homo sapien
110	41	91.1	190	2	Q7R5F8_GIALA	Q7r5f8 giardia lam	183	41	91.1	837	1	FA40A_MACFA	Q9glz5 macaca fasc
111	41	91.1	194	2	Q96B22_HUMAN	Q96b22 homo sapien	184	41	91.1	837	1	FA40A_MOUSE	Q8c079 mus musculus
112	41	91.1	197	2	Q7XF54_ORYSA	Q7xf54 oryza sativ	185	41	91.1	837	1	FA40A_PONPY	Q5r7a4 pongo pygma
113	41	91.1	197	2	Q8S780_ORYSA	Q8s780 oryza sativ	186	41	91.1	837	1	Q6NSB6_RHOPA	Q6nsb6 rhodopsaudo
114	41	91.1	205	2	Q4NRN0_9DELT	Q4nrn0 anaeromyxob	187	41	91.1	841	1	RELA_STRAT	Q85709 streptomyce
115	41	91.1	226	2	Q51TR6_MAGGR	Q51tr6 magnaporthie	188	41	91.1	851	2	Q55XY5_CRYNE	Q55xy5 cryptococcus
116	41	91.1	239	2	Q5SC69_DICDI	Q5sc69 dictyostelli	189	41	91.1	873	1	RX_DROME	Q9w2q1 drosophila
117	41	91.1	248	1	ICP34_HRV11	Q36313 human herpe	190	41	91.1	873	1	Q5KM08_CRYNE	Q5km08 cryptococcus
118	41	91.1	248	1	RL1_HRV11	Q12396 human herpe	191	41	91.1	874	2	Q5JLK9_ORYSA	Q5jlk9 oryza sativ
119	41	91.1	251	2	Q52IU9_MAGGR	Q52iu9 magnaporthie	192	41	91.1	883	2	Q529K7_MAGGR	Q529k7 magnaporthie
120	41	91.1	255	2	Q902R5_BRARE	Q902r5 brachydanio	193	41	91.1	895	2	Q7S902_NEUCR	Q7s902 neurospora
121	41	91.1	264	2	Q07J93_MYCAV	Q07j93 mycobacteri	194	41	91.1	896	2	Q96D18_HUMAN	Q96d18 homo sapien
122	41	91.1	273	2	Q96KW5_HUMAN	Q96kw5 homo sapien	195	41	91.1	902	2	Q68FL6_MOUSE	Q68fl6 mus musculus
123	41	91.1	277	2	Q8WYH1_HUMAN	Q8wyh1 homo sapien	196	41	91.1	972	2	Q9M2D4_ARATH	Q9m2d4 arabidopsis
124	41	91.1	279	2	Q6X235_9ALPH	Q6x235 bovine herp	197	41	91.1	979	1	RFK1_HUMAN	P22670 homo sapien
125	41	91.1	282	2	Q5COW6_SCHJA	Q5cow6 schistosoma	198	41	91.1	993	2	Q7VVD6_BORPE	Q7vvd6 bordetella
126	41	91.1	298	2	Q5M6G5_ORYSA	Q5m6g5 oryza sativ	199	41	91.1	1012	2	Q5VZ95_HUMAN	Q5vz95 homo sapien
127	41	91.1	317	2	Q8H4Q2_ORYSA	Q8h4q2 oryza sativ	200	41	91.1	1012	2	Q9UEF7_HUMAN	Q9uef7 homo sapien
128	41	91.1	320	1	HXA4_HUMAN	Q00056 homo sapien	201	41	91.1	1013	2	Q4SDR3_TETNG	Q4sdr3 tetraodon n
129	41	91.1	324	1	YBX2B_XENLA	P45441 xenopus lae	202	41	91.1	1013	2	Q8WP17_MACFA	Q8wp17 macaca fasc
130	41	91.1	327	2	Q8BHG7_MOUSE	Q8bhg7 m mus muscu	203	41	91.1	1029	2	Q9YMX0_NPVLID	Q9ymx0 lymnaetria d
131	41	91.1	338	2	Q6P0M8_HUMAN	Q6p0m8 homo sapien	204	41	91.1	1125	2	Q4RJL1_TETNG	Q4rjl1 tetraodon n
132	41	91.1	342	2	Q4RE66_TETNG	Q4re66 tetraodon n	205	41	91.1	1165	2	Q86HN1_DICDI	Q86hni dictyostell
133	41	91.1	343	2	Q6ZJE6_ORYSA	Q6zje6 oryza sativ	206	41	91.1	1165	2	Q4QHC6_LEIMA	Q4qhc6 leishmania
134	41	91.1	353	2	Q6Q502_ORYSA	Q6q502 oryza sativ	207	41	91.1	1188	2	Q7RW25_NEUCR	Q7rw25 neurospora
135	41	91.1	353	2	Q6Q7E3_ORYSA	Q6q7e3 oryza sativ	208	41	91.1	1217	2	Q9ULL5_HUMAN	Q9ull5 homo sapien
136	41	91.1	361	2	Q6ETY2_ORYSA	Q6ety2 oryza sativ	209	41	91.1	1229	2	Q68H99_PIG	Q68h99 sus scrofa
137	41	91.1	361	2	Q7IEP2_ORYSA	Q7iep2 oryza sativ	210	41	91.1	1251	1	IRSI_CERAE	Q28224 cercopithe
138	41	91.1	380	2	Q7X3Y7_9CARY	Q7x3y7 fagopyrum s	211	41	91.1	1253	2	Q6DHV6_HUMAN	Q6dhv6 homo sapien
139	41	91.1	381	2	Q7X3Y9_9CARY	Q7x3y9 fagopyrum g	212	41	91.1	1294	2	Q4SAM8_TETNG	Q4sam8 tetraodon n
140	41	91.1	381	2	Q7XXU0_ORYSA	Q7xxu0 oryza sativ	213	41	91.1	1338	1	IRS2_HUMAN	Q9y4h2 homo sapien
141	41	91.1	389	2	Q59NU2_CANAL	Q59nu2 candida alb	214	41	91.1	1338	2	Q96RG4_HUMAN	Q96rg4 homo sapien
142	41	91.1	390	2	Q59NP4_CANAL	Q59np4 candida alb	215	41	91.1	1339	2	Q96RG5_HUMAN	Q96rg5 homo sapien
143	41	91.1	418	1	NCOA6_RAT	Q9jli4 x nuclear r	216	41	91.1	1367	2	Q8WZV8_NEUCR	Q8wzv8 neurospora
144	41	91.1	427	1	BSH_DROME	Q04787 drosophila	217	41	91.1	1400	1	BRD4_MOUSE	Q9esue6 mus musculu
145	41	91.1	429	2	Q4V6V3_DROME	Q4v6v3 drosophila	218	41	91.1	1420	2	Q9Y0Y8_DROME	Q9y0y8 drosophila
146	41	91.1	451	2	Q7Q6A6_ANOGA	Q7q6a6 anopheles g	219	41	91.1	1430	2	Q9H7P9_HUMAN	Q9h7p9 homo sapien
147	41	91.1	456	2	Q7SH82_NEUCR	Q7sh82 neurospora	220	41	91.1	1506	2	Q8IRL0_DROME	Q8irl0 drosophila
148	41	91.1	458	2	Q4NSW5_9DELT	Q4nsw5 anaeromyxob	221	41	91.1	1641	2	Q5NCY0_MOUSE	Q5ncy0 mus musculu
149	41	91.1	482	1	PDCT7_MOUSE	Q9wty1 mus musculu	222	41	91.1	1641	2	Q4VC26_MOUSE	Q4vc26 mus musculu
150	41	91.1	485	1	PDCT7_HUMAN	Q8nsd1 homo sapien	223	41	91.1	1682	2	O15054_HUMAN	O15054 homo sapien
151	41	91.1	485	2	Q6IEG3_HUMAN	Q6ieg3 homo sapien	224	41	91.1	1998	1	MYST3_RAT	Q5tkr9 rattus norv
152	41	91.1	485	2	Q24136_DROME	Q24136 drosophila	225	41	91.1	2003	1	MYST3_MOUSE	Q8b221 mus musculu
153	41	91.1	485	2	Q9V552_DROME	Q9v552 drosophila	226	41	91.1	2004	1	MYST3_HUMAN	Q92794 h histone a
154	41	91.1	502	2	Q8NFC6_HUMAN	Q8nfc6 homo sapien	227	41	91.1	2133	2	Q8IQT4_DROME	Q8igt4 drosophila
155	41	91.1	520	2	Q4RQ72_TETNG	Q4rq72 tetraodon n	228	41	91.1	2134	2	Q4SJP0_TETNG	Q4sjp0 tetraodon n
156	41	91.1	551	2	Q8WP18_MACFA	Q8wp18 macaca fasc	229	41	91.1	2306	2	Q9WUH9_RAT	Q9wuh9 rattus norv
157	41	91.1	573	2	Q4S0L1_TETNG	Q4s0l1 tetraodon n	230	41	91.1	2307	1	FBN2_MOUSE	Q61555 mus musculu
158	41	91.1	590	2	Q7Y214_ARATH	Q7y214 arabidopsis	231	41	91.1	2311	1	FBN2_HUMAN	P35556 homo sapien
159	41	91.1	600	2	Q8NAJ6_HUMAN	Q8naj6 homo sapien	232	41	91.1	3144	1	HD_HUMAN	P42858 homo sapien
160	41	91.1	619	2	Q8UZE3_9GAMA	Q8uze3 cercopithe	233	41	91.1	3160	1	EP400_HUMAN	Q96191 homo sapien
161	41	91.1	621	2	Q7S8B6_NEUCR	Q7s8b6 neurospora	234	41	91.1	3469	2	Q9U4I2_DROME	Q9u4i2 drosophila
162	41	91.1	630	1	MECT1_MOUSE	Q68ed7 mus musculu	235	41	91.1	3571	2	Q4QIZ2_LEIMA	Q4qiq2 leishmania
163	41	91.1	634	1	MECT1_HUMAN	Q6uuu9 homo sapien	236	41	91.1	3604	2	Q9VYK0_DROME	Q9vyk0 drosophila
164	41	91.1	641	2	Q51ZC0_MAGGR	Q51zc0 magnaporthie	237	41	91.1	4911	1	MLL3_HUMAN	Q8nez4 homo sapien
165	41	91.1	643	2	Q9LSL2_ARATH	Q9lsl2 arabidopsis	238	41	91.1	43	2	P79301_PIG	P79301 sus scrofa
166	41	91.1	645	2	Q51V73_MAGGR	Q51v73 magnaporthie	239	41	91.1	44	2	Q8HXS5_CAPCA	Q8hxs5 capreolus c
167	41	91.1	660	2	Q61J16_CABER	Q61j16 caenorhabdi	240	41	91.1	51	2	Q8HXS4_CAPCA	Q8hxs4 capreolus c
168	41	91.1	684	2	Q7Y1D3_ORYSA	Q7y1d3 oryza sativ	241	41	91.1	52	2	Q8HXT2_CAPCA	Q8hxt2 capreolus c
169	41	91.1	689	2	Q8WPL5_9UROC	Q8wpl5 oikopleura	242	41	91.1	52	2	Q8HXT4_CAPCA	Q8hxt4 capreolus c
170	41	91.1	708	2	Q51LY7_MAGGR	Q51ly7 magnaporthie	243	41	91.1	67	2	Q6K657_ORYSA	Q6k657 oryza sativ
171	41	91.1	711	2	Q7XT28_ORYSA	Q7xt28 oryza sativ	244	41	91.1	70	2	Q9MAV4_ARATH	Q9mav4 arabidopsis
172	41	91.1	733	2	Q9VTX9_DROME	Q9vtx9 drosophila	245	41	91.1	89	2	Q6VTG3_ORYSA	Q6vtg3 oryza sativ
173	41	91.1	736	1	CTCF_MOUSE	Q61164 mus musculu	246	41	91.1	105	2	Q7S3N9_NEUCR	Q7s3n9 neurospora
174	41	91.1	766	2	Q4RK16_TETNG	Q4rk16 tetraodon n	247	41	91.1	109	2	Q40548_TOBAC	Q40548 nicotiana c
175	41	91.1	780	1	T22D2_HUMAN	Q75157 homo sapien	248	41	91.1	110	2	Q71DC1_DROYA	Q71dc1 drosophila
176	41	91.1	808	2	Q4P0G5_USTLGA	Q4p0g5 ustilago ma	249	41	91.1	110	2	Q9W154_DROME	Q9w154 drosophila
177	41	91.1	814	2	Q4IR23_GIBZE	Q4ir23 gibberella	250	41	91.1	115	2	Q5AL62_CANAL	Q5al62 candida alb

251	40	88.9	116	2	Q53SV4_HUMAN	Q53sv4 homo sapien	324	40	88.9	226	2	Q61H53_CAEBR	Q61h53 caenorhabdi
252	40	88.9	116	2	Q9VE17_DROME	Q9ve17 drosophila	325	40	88.9	227	2	Q66L28_XENIA	Q66l28 xenopus lae
253	40	88.9	118	2	Q9H387_HUMAN	Q9h387 homo sapien	326	40	88.9	231	2	Q4Q5X2_LEIMA	Q4q5x2 leishmania
254	40	88.9	118	2	Q9DIY5_MOUSE	Q9diy5 mus musculus	327	40	88.9	231	2	Q75IX1_ORYSA	Q75ix1 oryza sativ
255	40	88.9	121	2	Q4SP92_TETNG	Q4sp92 tetraodon n	328	40	88.9	235	2	Q8LILM9_ORYSA	Q8lilm9 oryza sativ
256	40	88.9	122	2	Q6ZNV9_homo sapien	Q6znv9 homo sapien	329	40	88.9	239	2	Q51HW0_MAGGR	Q51hw0 magnaporthe
257	40	88.9	128	2	Q4IE09_GIBZEA	Q4ie09 gibberella	330	40	88.9	240	2	P89268_GVXN	P89268 xestia c-ni
258	40	88.9	130	2	Q6ZSV2_HUMAN	Q6zsv2 homo sapien	331	40	88.9	242	2	Q43687_VIGUN	Q43687 vigna ungui
259	40	88.9	132	2	Q6ZNU7_HUMAN	Q6znu7 homo sapien	332	40	88.9	242	2	Q4RLI9_TETNG	Q4rll9 tetraodon n
260	40	88.9	135	2	Q95KE1_MACFA	Q95ke1 macaca fasc	333	40	88.9	243	2	Q84LD5_9ASTR	Q84ld5 chrysanthem
261	40	88.9	136	2	Q643A9_STRHY	Q643a9 streptomyce	334	40	88.9	246	2	Q6QE92_9MOLL	Q6qe92 pareledone
262	40	88.9	138	2	Q6ZUUS_HUMAN	Q6zuus homo sapien	335	40	88.9	247	2	Q75LM5_ORYSA	Q75lm5 oryza sativ
263	40	88.9	139	2	Q6Z571_ORYSA	Q6z571 oryza sativ	336	40	88.9	249	2	Q4US07_XANCP	Q4us07 xanthomonas
264	40	88.9	140	2	Q7XIT9_ORYSA	Q7xit9 oryza sativ	337	40	88.9	251	2	Q59RN8_CANAL	Q59rn8 candida alb
265	40	88.9	143	2	Q7PNB6_ANOGA	Q7pnb6 anopheles g	338	40	88.9	251	2	Q8LS83_ORYSA	Q8ls83 oryza sativ
266	40	88.9	147	2	Q5BMN9_SCHJA	Q5bm9 schistosoma	339	40	88.9	252	2	Q5CD49_BRARE	Q5cd49 brachydanio
267	40	88.9	148	2	Q6ZUC5_HUMAN	Q6zuc5 homo sapien	340	40	88.9	254	2	Q61WP8_CAEBR	Q61wp8 caenorhabdi
268	40	88.9	148	2	Q6ZMGO_HUMAN	Q6zmgo homo sapien	341	40	88.9	259	2	Q9ARE2_FLABI	Q9are2 flavaria bi
269	40	88.9	151	2	Q5VQU5_ORYSA	Q5vqu5 oryza sativ	342	40	88.9	261	2	Q94LQ0_ORYSA	Q94lq0 oryza sativ
270	40	88.9	152	2	Q6ZUG4_HUMAN	Q6zug4 homo sapien	343	40	88.9	265	2	Q6UX52_HUMAN	Q6ux52 homo sapien
271	40	88.9	153	2	Q6ERN8_ORYSA	Q6ern8 oryza sativ	344	40	88.9	265	2	Q5C4Q7_SCHJA	Q5c4q7 schistosoma
272	40	88.9	155	2	Q9HBV3_HUMAN	Q9hbv3 homo sapien	345	40	88.9	266	2	Q7XL73_ORYSA	Q7xl73 oryza sativ
273	40	88.9	156	2	Q5T481_HUMAN	Q5t481 homo sapien	346	40	88.9	267	2	Q6DW99_CHICK	Q6dw99 gallus gall
274	40	88.9	158	2	Q9DVM0_3BACU	Q9dvm0 plutella xy	347	40	88.9	275	2	Q9WIU7_DROME	Q9wiu7 drosophila
275	40	88.9	161	2	Q75LJ3_ORYSA	Q75lj3 oryza sativ	348	40	88.9	275	2	Q7YU61_DROME	Q7yu61 drosophila
276	40	88.9	162	2	Q7XAL2_ORYSA	Q7xal2 oryza sativ	349	40	88.9	277	2	Q6K296_ORYSA	Q6k296 oryza sativ
277	40	88.9	162	2	Q9AMM3_ORYSA	Q9amm3 oryza sativ	350	40	88.9	280	2	Q86FT3_9TURB	Q86ft3 dugesia str
278	40	88.9	165	2	Q77LW4_9NUCL	Q77lw4 heliocoverpa	351	40	88.9	282	2	Q6WP44_MUSSP	Q6wp44 mus spretus
279	40	88.9	165	2	Q77LZ7_9NUCL	Q77lz7 heliocoverpa	352	40	88.9	282	2	Q6WP45_MUSSP	Q6wp45 mus spretus
280	40	88.9	165	2	Q8V5R8_9NUCL	Q8v5r8 heliocoverpa	353	40	88.9	282	2	Q6WP46_MUSSP	Q6wp46 mus spretus
281	40	88.9	165	2	Q91BW0_9NUCL	Q91bw0 heliocoverpa	354	40	88.9	282	2	Q6WP48_MUSSP	Q6wp48 mus spretus
282	40	88.9	165	2	Q99GW5_9NUCL	Q99gw5 heliocoverpa	355	40	88.9	283	2	Q5XS8_ARATH	Q5xs8 arabidopsis
283	40	88.9	171	2	Q7DLN2_TOBAC	Q7dln2 nicotiana t	356	40	88.9	283	2	Q6ZID9_ORYSA	Q6zid9 oryza sativ
284	40	88.9	177	2	Q7XQ64_ORYSA	Q7xq64 oryza sativ	357	40	88.9	283	2	Q6Z6Q4_ORYSA	Q6z6q4 oryza sativ
285	40	88.9	177	2	Q8H062_ORYSA	Q8h062 oryza sativ	358	40	88.9	291	2	Q6BH87_DEBHA	Q6bh87 debaryomyce
286	40	88.9	178	2	Q69T06_ORYSA	Q69t06 oryza sativ	359	40	88.9	296	2	Q7Q6Q5_ANOGA	Q7q6q5 anopheles g
287	40	88.9	180	2	Q8NAQ6_HUMAN	Q8naq6 homo sapien	360	40	88.9	297	2	Q9JLV7_MOUSE	Q9jlv7 mus musculus
288	40	88.9	182	2	Q7XXM6_ORYSA	Q7xxm6 oryza sativ	361	40	88.9	298	2	Q5QNA1_ORYSA	Q5qn1 oryza sativ
289	40	88.9	184	2	Q5VRR2_ORYSA	Q5vrr2 oryza sativ	362	40	88.9	302	2	Q9C889_ARATH	Q9c889 arabidopsis
290	40	88.9	185	2	Q5QLQ3_ORYSA	Q5qlq3 oryza sativ	363	40	88.9	303	2	Q9C8B9_ARATH	Q9c8b9 arabidopsis
291	40	88.9	191	2	Q51ZW1_MAGGR	Q51zw1 magnaporthe	364	40	88.9	302	2	Q84711_9LUTE	Q84711 pea enation
292	40	88.9	192	2	Q7S944_NEUCR	Q7s944 neurospora	365	40	88.9	306	2	Q94HB0_ORYSA	Q94hb0 oryza sativ
293	40	88.9	196	2	Q81955_TOBAC	Q81955 nicotiana t	366	40	88.9	310	2	Q527E0_MAGGR	Q527e0 magnaporthe
294	40	88.9	196	2	Q84UR6_ORYSA	Q84ur6 oryza sativ	367	40	88.9	310	2	Q5ZDE8_ORYSA	Q5zde8 oryza sativ
295	40	88.9	197	2	Q413A7_GIBZEA	Q413a7 gibberella	368	40	88.9	312	2	Q4NL12_9MICC	Q4nl12 arthrobacte
296	40	88.9	198	2	Q417M8_GIBZEA	Q417m8 gibberella	369	40	88.9	313	1	CDX2_HUMAN	CDX22 human
297	40	88.9	203	2	Q4QBL7_LEIMA	Q4qbl7 leishmania	370	40	88.9	313	1	ENV1_MOUSE	ENV1 mouse
298	40	88.9	205	2	Q69W83_ORYSA	Q69w83 oryza sativ	371	40	88.9	313	2	Q5VTU7_HUMAN	Q5vtu7 homo sapien
299	40	88.9	206	2	Q82522_STRAW	Q82522 streptomyce	372	40	88.9	315	2	Q5ISK4_MACFA	Q5isk4 macaca fasc
300	40	88.9	207	2	Q8EFY6_SHEON	Q8efy6 shewanella	373	40	88.9	316	2	Q57XT4_9TRYP	Q57xt4 trypanosoma
301	40	88.9	209	2	Q8L194_TOBAC	Q8l194 nicotiana t	374	40	88.9	317	2	Q6PB42_MOUSE	Q6pb42 mus musculus
302	40	88.9	212	2	Q7Y1G1_ORYSA	Q7y1g1 oryza sativ	375	40	88.9	319	2	Q9FLU5_ARATH	Q9flu5 arabidopsis
303	40	88.9	213	2	Q6SQ84_HDV	Q6sq84 hepatitis d	376	40	88.9	320	2	Q6ZSA5_HUMAN	Q6zsa5 homo sapien
304	40	88.9	213	2	Q760D0_HDV	Q760d0 hepatitis d	377	40	88.9	320	2	Q8L9K2_ARATH	Q8l9k2 arabidopsis
305	40	88.9	213	2	Q760D2_HDV	Q760d2 hepatitis d	378	40	88.9	322	2	Q4NU48_9DELT	Q4nu48 anaeromyxob
306	40	88.9	213	2	Q760D3_HDV	Q760d3 hepatitis d	379	40	88.9	324	2	Q8LHR2_ORYSA	Q8lhr2 oryza sativ
307	40	88.9	213	2	Q760D4_HDV	Q760d4 hepatitis d	380	40	88.9	325	2	Q9BSN0_HUMAN	Q9bsn0 homo sapien
308	40	88.9	213	2	Q760D5_HDV	Q760d5 hepatitis d	381	40	88.9	325	2	Q4RTH8_TETNG	Q4rth8 tetraodon n
309	40	88.9	213	2	Q9PWX0_HDV	Q9pwx0 hepatitis d	382	40	88.9	326	2	Q91WA8_MOUSE	Q91wa8 mus musculus
310	40	88.9	213	2	Q9PX07_HDV	Q9px07 hepatitis d	383	40	88.9	328	2	Q5VVK1_ORYSA	Q5vnk1 oryza sativ
311	40	88.9	213	2	Q9PZV7_HDV	Q9pzv7 hepatitis d	384	40	88.9	328	2	Q64262_MOUSE	Q64262 mus musculus
312	40	88.9	213	2	Q9PZV8_HDV	Q9pzv8 hepatitis d	385	40	88.9	329	2	Q8H8U0_ORYSA	Q8h8u0 oryza sativ
313	40	88.9	213	2	Q9PZV9_HDV	Q9pzv9 hepatitis d	386	40	88.9	334	2	Q7TZB1_MYCBO	Q7tzb1 mycobacteri
314	40	88.9	213	2	Q9PZW0_HDV	Q9pzw0 hepatitis d	387	40	88.9	334	2	Q07721_MYCTU	Q07721 mycobacteri
315	40	88.9	213	2	Q9PZW1_HDV	Q9pzw1 hepatitis d	388	40	88.9	335	2	Q6H830_ORYSA	Q6h830 oryza sativ
316	40	88.9	213	2	Q9PZW2_HDV	Q9pzw2 hepatitis d	389	40	88.9	340	2	Q922N8_MOUSE	Q922n8 mus musculus
317	40	88.9	213	2	Q9PZW3_HDV	Q9pzw3 hepatitis d	390	40	88.9	340	2	Q6KAU4_MOUSE	Q6kau4 mus musculus
318	40	88.9	215	2	Q61685_MOUSE	Q61685 mus musculus	391	40	88.9	342	2	Q5VPC0_ORYSA	Q5vpc0 oryza sativ
319	40	88.9	217	2	Q7XU26_ORYSA	Q7xu26 oryza sativ	392	40	88.9	343	2	Q8LNI0_ORYSA	Q8lni0 oryza sativ
320	40	88.9	221	2	Q8N657_HUMAN	Q8n657 homo sapien	393	40	88.9	344	1	TILS_SYNEL	TILS synelecoc
321	40	88.9	221	2	Q49946_SOLTU	Q49946 solanum tub	394	40	88.9	344	2	Q9H7E6_HUMAN	Q9h7e6 homo sapien
322	40	88.9	222	2	Q95S23_DROME	Q95s23 drosophila	395	40	88.9	345	2	Q7Q4G5_ANOGA	Q7q4g5 anopheles g
323	40	88.9	222	2	Q4TA38_TETNG	Q4ta38 tetraodon n	396	40	88.9	346	1	E2F5_HUMAN	E2f529 homo sapien

397	40	88.9	349	2	Q9P555_NEUCR	470	40	88.9	444	2	Q5AGS1_CANAL	Q5ag1 candida alb
398	40	88.9	350	1	HNFG3_HUMAN	471	40	88.9	444	2	Q5MNP0_9GAMR	Q5mnp0 murine leuk
399	40	88.9	350	2	Q53F16_HUMAN	472	40	88.9	445	2	Q5MNN8_9GAMR	Q5mnn8 murine leuk
400	40	88.9	351	1	KLF2_RAT	473	40	88.9	451	2	Q969F2_HUMAN	Q969f2 homo sapien
401	40	88.9	352	2	Q61SE7_HUMAN	474	40	88.9	451	2	Q4U2V9_CHLRE	Q4u2v9 chlamydomon
402	40	88.9	353	2	Q6TRW9_CHICK	475	40	88.9	453	2	Q96JEU9_HUMAN	Q96jEU9 homo sapien
403	40	88.9	354	1	ATOH1_HUMAN	476	40	88.9	454	2	Q69022_9GAMA	Q69022 human herpe
404	40	88.9	354	2	Q55F03_DICDI	477	40	88.9	455	2	Q90120_9GAMR	Q90120 mink cell f
405	40	88.9	354	2	Q41J14_GIBZE	478	40	88.9	456	2	Q8IRW4_DROME	Q8irw4 drosophila
406	40	88.9	355	2	Q478C9_TETNG	479	40	88.9	457	1	CHKA_HUMAN	CHKA_HUMAN
407	40	88.9	355	2	Q478C9_TETNG	480	40	88.9	457	2	Q4PIE9_USTWA	Q4pie9 ustiliago ma
408	40	88.9	356	1	ENV_FANTR	481	40	88.9	463	2	Q6DHH6_BRARE	Q6dhh6 brachydanio
409	40	88.9	356	2	Q6TRX0_CHICK	482	40	88.9	469	1	FOXGA_HUMAN	FOXGA_HUMAN
410	40	88.9	356	2	Q6TRX0_CHICK	483	40	88.9	474	2	Q35102_RATRT	Q35102 rattus ratt
411	40	88.9	356	2	Q87044_9GAMR	484	40	88.9	477	1	FOXGB_HUMAN	FOXGB_HUMAN
412	40	88.9	360	2	Q4QDS7_LEIMA	485	40	88.9	477	2	Q851M1_ORYSA	Q851m1 oryza sativ
413	40	88.9	361	2	Q93VV8_ORYSA	486	40	88.9	477	2	Q68800_FRMCV	Q68800 friend mink
414	40	88.9	369	2	Q61Z90_CAEBR	487	40	88.9	480	1	FOXGB_RAT	FOXGB_RAT
415	40	88.9	372	2	Q60TD1_CAEBR	488	40	88.9	481	1	FOXGB_MOUSE	FOXGB_MOUSE
416	40	88.9	374	2	Q525P5_MAGGR	489	40	88.9	485	2	Q7XB19_ORYSA	Q7xb19 oryza sativ
417	40	88.9	374	2	Q5U044_ANOGA	490	40	88.9	486	2	Q8LAL7_ARATH	Q8lal7 arabidopsis
418	40	88.9	376	2	Q4XOM9_ASFPV	491	40	88.9	486	2	Q9FNN5_HUMAN	Q9fnn5 arabidopsis
419	40	88.9	376	2	Q4SD35_TETNG	492	40	88.9	489	2	Q86XT7_HUMAN	Q86xt7 homo sapien
420	40	88.9	377	2	Q6GUH1_HEVER	493	40	88.9	489	2	Q8K557_MOUSE	Q8k557 mus musculus
421	40	88.9	378	2	Q6T925_CHICK	494	40	88.9	490	2	Q6PDH9_MOUSE	Q6pdh9 mus musculus
422	40	88.9	379	2	Q504Y4_HUMAN	495	40	88.9	490	2	Q69JG3_ORYSA	Q69jg3 oryza sativ
423	40	88.9	387	2	Q82LR9_STRAW	496	40	88.9	491	2	Q69JG3_ORYSA	Q69jg3 oryza sativ
424	40	88.9	388	2	Q8MSD8_DROME	497	40	88.9	493	2	Q5BHV6_DROSOPHILA	Q5bhv6 drosophila
425	40	88.9	388	2	Q84T11_ORYSA	498	40	88.9	493	2	Q09708_PEAUV	Q09708 pea enation
426	40	88.9	389	2	Q73TN8_MYCPA	499	40	88.9	493	2	Q93184_PEAUV	Q93184 pea enation
427	40	88.9	390	1	DAF1_MOUSE	500	40	88.9	495	1	PO3F3_MOUSE	PO3F3_MOUSE
428	40	88.9	390	1	OST3B_HUMAN	501	40	88.9	497	2	Q5QEU3_9ACTO	Q5qeu3 streptomyc
429	40	88.9	390	1	OST3B_MOUSE	502	40	88.9	498	2	Q688S1_9AGAR	Q688s1 pleurotomy
430	40	88.9	390	2	Q5NCR0_MOUSE	503	40	88.9	500	1	PO3F3_HUMAN	PO3F3_HUMAN
431	40	88.9	390	2	Q921P0_MOUSE	504	40	88.9	500	2	Q4ZG25_HUMAN	Q4zg25 homo sapien
432	40	88.9	390	2	Q4FJ84_MOUSE	505	40	88.9	500	2	Q5QLD5_ORYSA	Q5qlD5 oryza sativ
433	40	88.9	391	2	Q54M74_DICDI	506	40	88.9	501	2	Q4RWZ6_TETNG	Q4rwz6 tetraodon n
434	40	88.9	394	2	Q82249_PYRAE	507	40	88.9	503	2	Q8W8U5_9METZ	Q8w8u5 ephydatia f
435	40	88.9	396	2	Q522C4_MAGGR	508	40	88.9	504	2	Q8H2T7_ORYSA	Q8h2t7 oryza sativ
436	40	88.9	397	2	Q85733_FRSPV	509	40	88.9	505	2	Q4P5S4_USTMA	Q4p5s4 ustiliago ma
437	40	88.9	402	2	Q9UPV1_HUMAN	510	40	88.9	506	1	WRK10_ARATH	WRK10_ARATH
438	40	88.9	403	2	Q5JN13_ORYSA	511	40	88.9	509	2	Q5MNN2_9GAMR	Q5mnn2 murine leuk
439	40	88.9	406	2	Q6H4X7_ORYSA	512	40	88.9	520	2	Q5VR46_ORYSA	Q5vr46 oryza sativ
440	40	88.9	408	1	ENV_RSFFV	513	40	88.9	526	2	Q7SEK3_NEUCR	Q7sek3 neurospora
441	40	88.9	408	1	INHEB_BOVIN	514	40	88.9	527	2	Q54ND4_DICDI	Q54nd4 dictyostell
442	40	88.9	410	2	Q61K70_CAEBR	515	40	88.9	529	2	Q5MNM0_9GAMR	Q5mnm0 murine leuk
443	40	88.9	410	2	Q7PXR9_ANOGA	516	40	88.9	532	2	Q6PK2_ORYSA	Q6pk2 oryza sativ
444	40	88.9	410	2	Q92791_FRSPV	517	40	88.9	533	1	CDPK2_ORYSA	CDPK2_ORYSA
445	40	88.9	411	2	Q6OR98_CAEBR	518	40	88.9	533	2	Q8GTY8_ORYSA	Q8gty8 oryza sativ
446	40	88.9	412	1	ALP_PETHY	519	40	88.9	534	1	PTZF1_BOMMO	PTZF1_BOMMO
447	40	88.9	412	2	Q651H6_ORYSA	520	40	88.9	535	2	Q5MNL0_9GAMR	Q5mnl0 murine leuk
448	40	88.9	413	2	Q9SVU6_ARATH	521	40	88.9	535	2	Q5MNL8_9GAMR	Q5mnl8 murine leuk
449	40	88.9	415	2	Q9VT88_DROME	522	40	88.9	535	2	Q5MNM2_9GAMR	Q5mnm2 murine leuk
450	40	88.9	416	1	HXD3_HUMAN	523	40	88.9	535	2	Q5MNN2_9GAMR	Q5mnn2 murine leuk
451	40	88.9	416	2	Q95TV8_DROME	524	40	88.9	535	2	Q5MNN4_9GAMR	Q5mnn4 murine leuk
452	40	88.9	417	1	HXD3_MOUSE	525	40	88.9	537	2	Q55G00_DICDI	Q55g00 dictyostell
453	40	88.9	417	2	Q8SX17_DROME	526	40	88.9	537	2	Q4NTG0_9DEL	Q4ntg0 anaeromyxob
454	40	88.9	422	2	Q9HXT5_PSEAE	527	40	88.9	546	2	Q86XP8_HUMAN	Q86xp8 homo sapien
455	40	88.9	425	2	Q6YX84_ORYSA	528	40	88.9	546	2	Q40451_TOBAC	Q40451 nicotiana t
456	40	88.9	425	2	Q8RUJ8_ORYSA	529	40	88.9	549	2	Q61707_CABEL	Q61707 caenorhabdi
457	40	88.9	427	2	Q64ON2_MOUSE	530	40	88.9	553	2	Q9W153_DROME	Q9w153 drosophila
458	40	88.9	428	1	FOXZ2_MOUSE	531	40	88.9	553	2	Q91P58_MYXXA	Q91p58 myxococcus
459	40	88.9	431	1	KXAT3_ARATH	532	40	88.9	556	2	Q6B1Q2_DEBHA	Q6b1q2 debaryomyce
460	40	88.9	431	2	Q8LDC1_ARATH	533	40	88.9	557	1	WRT4_CABEL	WRT4_CABEL
461	40	88.9	432	2	Q96HM7_HUMAN	534	40	88.9	558	2	Q96D88_HUMAN	Q96d88 homo sapien
462	40	88.9	432	2	Q5VYV0_HUMAN	535	40	88.9	559	2	Q943J0_ORYSA	Q943j0 oryza sativ
463	40	88.9	435	2	Q5MNN0_9GAMR	536	40	88.9	561	2	Q9H2M8_HUMAN	Q9h2m8 homo sapien
464	40	88.9	436	2	Q7SBR4_NEUCR	537	40	88.9	562	2	Q8NAP0_HUMAN	Q8naf0 homo sapien
465	40	88.9	439	2	Q6X286_RETFL	538	40	88.9	562	2	Q54EY4_DICDI	Q54ey4 dictyostell
466	40	88.9	439	2	Q6X286_RETFL	539	40	88.9	562	2	Q916Y0_MYXXA	Q916y0 myxococcus
467	40	88.9	440	2	Q01672_ASCSU	540	40	88.9	562	2	Q80VM4_MOUSE	Q80vm4 mus musculus
468	40	88.9	440	2	Q85561_9GAMR	541	40	88.9	564	2	Q7KZY0_HUMAN	Q7kzy0 homo sapien
469	40	88.9	442	2	Q8N7F2_HUMAN	542	40	88.9	568	2	Q55J33_CRYNE	Q55j33 cryptococcus

543	40	88.9	568	2	05KCM3_CRYNE	05kcm3 cryptococcus	616	40	88.9	645	2	09H9W9_HUMAN	Q9h9w9 homo sapien
544	40	88.9	572	2	08GZV8_ORYSA	08gzv8 oryza sativ	617	40	88.9	646	2	04HY89_GIBZE	04hy89 gibberella
545	40	88.9	576	2	05ZK52_CHICK	05zk52 gallus gall	618	40	88.9	653	2	073TB8_MYCPA	073tb8 mycobacteri
546	40	88.9	579	1	08WSJ3_MOUSE	08wsh3 mus musculus	619	40	88.9	653	2	0810W5_SPEPR	0810w5 sperophilu
547	40	88.9	582	2	08WSJ3_MANSE	08wsh3 manduca sex	620	40	88.9	655	1	FOX01_HUMAN	Q12778 homo sapien
548	40	88.9	582	2	0632X2_MOUSE	0632x2 mus musculus	621	40	88.9	655	2	05VYC7_HUMAN	05vyc7 homo sapien
549	40	88.9	585	2	08WTX2_HUMAN	08wtx2 homo sapien	622	40	88.9	655	2	08YMK1_ANASP	Q8ymk1 anabena sp
550	40	88.9	585	2	09BWN1_HUMAN	09bwn1 homo sapien	623	40	88.9	658	2	08MQZ8_DROME	08mqz8 drosophila
551	40	88.9	586	2	094GR7_ORYSA	094gr7 oryza sativ	624	40	88.9	659	1	MMPL5_HUMAN	P15111 homo sapien
552	40	88.9	586	2	058EN3_BRARE	058en3 brachydanio	625	40	88.9	669	2	054E44_DICDI	054e44 dictyosteli
553	40	88.9	588	2	05AM01_CANAL	05am01 candida alb	626	40	88.9	678	2	0870Q4_NEUCR	0870q4 neurospora
554	40	88.9	588	2	094BY8_VOLCA	094by8 volvox cart	627	40	88.9	683	2	08LIW1_ORYSA	Q8liw1 oryza sativ
555	40	88.9	588	2	09SNE9_ARATH	09sne9 arabidopsis	628	40	88.9	692	2	04QA44_LEIMA	04qa44 leishmania
556	40	88.9	589	2	05ALK4_CANAL	05alk4 candida alb	629	40	88.9	693	2	04T7R3_TETNG	Q4t7r3 tetraodon n
557	40	88.9	590	2	061876_MOUSE	061876 mus musculus	630	40	88.9	696	2	060YB5_CABER	060yb5 caenorhabdi
558	40	88.9	593	1	ALUE_HUMAN	P39193 homo sapien	631	40	88.9	701	2	07Z287_APLCA	Q7z287 aplysia cal
559	40	88.9	594	2	09VEP4_DROME	09vep4 drosophila	632	40	88.9	706	2	09VNZ5_DROME	Q9vnz5 drosophila
560	40	88.9	597	2	055VE2_CRYNE	055ve2 cryptococcus	633	40	88.9	707	2	05NEU0_ORYSA	Q5neu0 oryza sativ
561	40	88.9	597	2	05KKU4_CRYNE	05kku4 cryptococcus	634	40	88.9	708	2	09LJG1_ARATH	Q9ljg1 arabidopsis
562	40	88.9	600	2	07YXU4_DICDI	07yxu4 dictyosteli	635	40	88.9	710	2	09PYW6_GVXN	Q9pyw6 xestia c-ni
563	40	88.9	601	2	07S9Q4_NEUCR	07s9q4 neurospora	636	40	88.9	712	1	ZBT39_HUMAN	015060 homo sapien
564	40	88.9	602	2	04G3H3_BRARE	04g3h3 brachydanio	637	40	88.9	714	2	Q4I754_GIBZE	Q4i754 gibberella
565	40	88.9	604	2	05R6Z2_PONPY	05r6z2 pongo pygma	638	40	88.9	717	2	05AMH8_CANAL	05amh8 candida alb
566	40	88.9	604	2	0651Z0_ORYSA	0651z0 oryza sativ	639	40	88.9	718	2	Q4TA13_TETNG	Q4ta13 tetraodon n
567	40	88.9	605	1	NCKX4_HUMAN	Q8nfz2 homo sapien	640	40	88.9	720	2	Q4I4D0_GIBZE	Q4i4d0 gibberella
568	40	88.9	605	1	NCKX4_MOUSE	Q8c9f8 mus musculus	641	40	88.9	721	2	Q8N8Y6_HUMAN	Q8n8y6 homo sapien
569	40	88.9	606	2	0852P0_VOLCA	Q852p0 volvox cart	642	40	88.9	721	2	Q8N3Y3_HUMAN	Q8n3y3 homo sapien
570	40	88.9	609	2	Q4RH16_TETNG	Q4rh16 tetraodon n	643	40	88.9	734	2	Q869R9_DICDI	Q869r9 dictyosteli
571	40	88.9	610	2	Q7QA42_ANOGA	Q7qa42 anopheles g	644	40	88.9	737	2	Q5NBD4_ORYSA	Q5nbd4 oryza sativ
572	40	88.9	610	2	066650_9GAMA	Q66650 equid herpe	645	40	88.9	740	2	Q59U10_CANAL	Q59u10 candida alb
573	40	88.9	612	2	Q7TPN9_MOUSE	Q7tpn9 mus musculus	646	40	88.9	742	2	Q59U34_CANAL	Q59u34 candida alb
574	40	88.9	613	2	Q6Z697_ORYSA	Q6z697 oryza sativ	647	40	88.9	744	2	Q4RMG0_TETNG	Q4rmg0 tetraodon n
575	40	88.9	615	2	Q9UDW8_HUMAN	Q9udw8 homo sapien	648	40	88.9	751	1	PASTA_MOUSE	Q8bi7 mus musculus
576	40	88.9	615	2	Q7PIF2_ANOGA	Q7pif2 anopheles g	649	40	88.9	751	1	PASTA_RAT	Q8443 rattus norv
577	40	88.9	616	1	VGF_HUMAN	Q15240 homo sapien	650	40	88.9	751	2	Q566E3_RAT	Q566e3 rattus norv
578	40	88.9	616	2	Q8T0J3_DROME	Q8t0j3 drosophila	651	40	88.9	756	2	018284_CAEBL	018284 caenorhabdi
579	40	88.9	617	1	VGF_RAT	P20156 rattus norv	652	40	88.9	767	2	Q7ZVZ3_BRARE	Q7zvz3 brachydanio
580	40	88.9	619	2	Q6H672_ORYSA	Q6h672 oryza sativ	653	40	88.9	768	2	Q5AFN4_CANAL	Q5afn4 candida alb
581	40	88.9	622	1	LAMO_DROME	Q08928 drosophila	654	40	88.9	768	2	Q4WZK4_ASPFU	Q4wzk4 aspergillus
582	40	88.9	622	2	Q9VMQ0_DROME	Q9vmq0 drosophila	655	40	88.9	773	2	Q9UJ19_DROME	Q9uj19 drosophila
583	40	88.9	625	2	Q52F37_MAGGR	Q52f37 magnaporth	656	40	88.9	775	2	Q8RW99_ARATH	Q8rw99 arabidopsis
584	40	88.9	625	2	Q7QC98_ANOGA	Q7qc98 anopheles g	657	40	88.9	781	2	Q4SB67_TETNG	Q4sb67 tetraodon n
585	40	88.9	628	2	Q5SHU0_THET8	Q5shu0 thermus the	658	40	88.9	782	1	PASTA_HUMAN	Q9y2w3 homo sapien
586	40	88.9	630	2	Q59VE3_CANAL	Q59ve3 candida alb	659	40	88.9	785	2	P74189_SYNY3	P74189 synechoyst
587	40	88.9	630	2	Q59V98_CANAL	Q59v98 candida alb	660	40	88.9	786	2	Q4HUC8_GIBZE	Q4huc8 gibberella
588	40	88.9	630	2	Q72I55_THET2	Q72i55 thermus the	661	40	88.9	787	2	Q9VEJ1_DROME	Q9vej1 drosophila
589	40	88.9	631	2	Q6EUM2_XENLA	Q6eum2 xenopus lae	662	40	88.9	791	2	Q5ZM84_CHICK	Q5zm84 gallus gall
590	40	88.9	632	2	Q6YXW2_ORYSA	Q6yxw2 oryza sativ	663	40	88.9	792	2	Q653C0_ORYSA	Q653c0 oryza sativ
591	40	88.9	633	2	Q7X188_ORYSA	Q7x188 oryza sativ	664	40	88.9	792	2	Q6EUD5_ORYSA	Q6eud5 oryza sativ
592	40	88.9	635	2	Q9WHJ7_FRMCV	Q9whj7 friend mink	665	40	88.9	795	2	Q504V9_HUMAN	Q504v9 homo sapien
593	40	88.9	636	1	ENV_MCF3	P15073 mink cell f	666	40	88.9	802	2	Q6GPA9_XENLA	Q6gpa9 xenopus lae
594	40	88.9	636	2	Q6WFA0_9GAMR	Q6wfa0 murine leuk	667	40	88.9	804	2	Q9UPS6_HUMAN	Q9ups6 homo sapien
595	40	88.9	636	2	Q6WPA1_9GAMR	Q6wpa1 murine leuk	668	40	88.9	804	2	Q4RX49_TETNG	Q4rx49 tetraodon n
596	40	88.9	636	2	Q6WPA2_9GAMR	Q6wpa2 murine leuk	669	40	88.9	805	2	Q61PW9_CABER	Q61pw9 caenorhabdi
597	40	88.9	636	2	Q6WPA3_9GAMR	Q6wpa3 murine leuk	670	40	88.9	812	2	Q8LJ56_ORYSA	Q8lj56 oryza sativ
598	40	88.9	636	2	Q85506_9GAMR	Q85506 murine leuk	671	40	88.9	815	2	Q61EP1_CABER	Q61ep1 caenorhabdi
599	40	88.9	638	2	Q9DQ21_9GAMR	Q9dq21 murine leuk	672	40	88.9	826	1	LOZEN_DROME	Q9w349 drosophila
600	40	88.9	639	2	Q6T3B1_9GAMR	Q6t3b1 murine leuk	673	40	88.9	831	2	Q7XPI3_ORYSA	Q7xpi3 oryza sativ
601	40	88.9	640	1	ENV_MCF3	P03388 mink cell f	674	40	88.9	834	2	Q5ACZ0_CANAL	Q5acz0 candida alb
602	40	88.9	640	1	ENV_RMCFV	P06445 rauscher mi	675	40	88.9	835	2	Q6PGC9_MOUSE	Q6pgc9 mus musculu
603	40	88.9	640	1	Q9HRV5_9GAMR	Q9hrv5 murine leuk	676	40	88.9	839	2	Q9SN46_ARATH	Q9sn46 arabidopsis
604	40	88.9	641	2	Q6WPA7_MUSGP	Q6wpa7 mus spretus	677	40	88.9	840	2	Q6U7W5_9TRYP	Q6u7w5 trypanosoma
605	40	88.9	641	2	Q78N71_MOUSE	Q78n71 mus musculus	678	40	88.9	847	2	Q6AU77_ORYSA	Q6au77 oryza sativ
606	40	88.9	641	2	Q78N73_MOUSE	Q78n73 mus musculus	679	40	88.9	849	2	Q520N2_MAGGR	Q520n2 magnaporth
607	40	88.9	641	2	Q78N94_MOUSE	Q78n94 mus musculus	680	40	88.9	850	1	SEMLA_DROME	Q24322 drosophila
608	40	88.9	641	2	Q80SW7_MOUSE	Q80sw7 mus musculus	681	40	88.9	851	2	Q4POG7_USTWA	Q4pog7 ustilago ma
609	40	88.9	641	2	Q80SW8_MOUSE	Q80sw8 mus musculus	682	40	88.9	854	2	Q4WFS5_ASPFU	Q4wfs5 aspergillus
610	40	88.9	641	2	Q80SW9_MOUSE	Q80sw9 mus musculus	683	40	88.9	854	2	Q5JXM8_HUMAN	Q5jxm8 homo sapien
611	40	88.9	641	2	Q80SY1_MOUSE	Q80sy1 mus musculus	684	40	88.9	855	2	Q80TK9_MOUSE	Q80tk9 mus musculu
612	40	88.9	641	2	Q81LM9_MOUSE	Q81lm9 mus musculus	685	40	88.9	856	2	Q6ZQD1_MOUSE	Q6zqd1 mus musculu
613	40	88.9	641	2	Q81LN4_MOUSE	Q81ln4 mus musculus	686	40	88.9	859	2	Q4QG27_LEINA	Q4qg27 leishmania
614	40	88.9	641	2	Q83363_9GAMR	Q83363 murine leuk	687	40	88.9	859	2	Q48495_TETNG	Q48495 tetraodon n
615	40	88.9	644	2	Q5RKN8_MOUSE	Q5rkn8 mus musculus	688	40	88.9	875	2	Q48930_SOYEN	Q48930 glycine max

689	40	88.9	890	2	Q9SVR0_SOYBN	Q9syr0 glycine max	762	40	88.9	1239	2	Q9V4P1_DROME	Q9v4p1 drosophila
690	40	88.9	898	2	Q4IMV1_GIBZE	Q4imv1 gibberella	763	40	88.9	1240	2	P79773_CHICK	P79773 gallus gall
691	40	88.9	899	2	Q8N304_HUMAN	Q8n304 homo sapien	764	40	88.9	1242	1	IRS1_HUMAN	IRS1 homo sapien
692	40	88.9	905	2	Q7XPUB_ORYSA	Q7xpub oryza sativ	765	40	88.9	1267	2	Q98945_CHICK	Q98945 gallus gall
693	40	88.9	911	2	Q8CGD1_MOUSE	Q8cgd1 mus musculu	766	40	88.9	1283	2	Q8T5H0_ANOGA	Q8t5h0 anopheles g
694	40	88.9	917	2	Q8CFQ8_MOUSE	Q8cfq8 mus musculu	767	40	88.9	1316	2	Q7S128_HUMAN	Q7s128 homo sapien
695	40	88.9	919	2	Q6NZLO_MOUSE	Q6nzlo mus musculu	768	40	88.9	1351	1	OVO_DROME	O51521 drosophila
696	40	88.9	932	2	Q54SG9_DICDI	Q54sg9 dictyosteli	769	40	88.9	1374	2	Q6AWI5_DROME	Q6aw15 drosophila
697	40	88.9	938	2	Q96VI4_PNECA	Q96vi4 pneumocysti	770	40	88.9	1388	2	Q9BYU8_HUMAN	Q9byu8 homo sapien
698	40	88.9	947	2	Q5TR21_HUMAN	Q5tf21 homo sapien	771	40	88.9	1396	2	Q4QJB8_LEIMA	Q4qjb8 leishmania
699	40	88.9	950	2	Q4P2X1_USTWA	Q4p2x1 ustilago ma	772	40	88.9	1403	1	PROS_DROME	P29617 drosophila
700	40	88.9	954	2	Q60TB3_CAEBR	Q60tb3 caenorhabdi	773	40	88.9	1425	2	Q4P427_USTMA	Q4p427 ustilago ma
701	40	88.9	961	2	Q51763_MAGGR	Q51t63 magnaporthie	774	40	88.9	1426	2	O15837_LEIMA	O15837 leishmania
702	40	88.9	963	1	RFX1_MOUSE	P48377 mus musculu	775	40	88.9	1437	2	Q9BYU9_HUMAN	Q9byu9 homo sapien
703	40	88.9	963	2	Q54F88_DICDI	Q54f88 dictyosteli	776	40	88.9	1437	2	Q9BZ95_HUMAN	Q9bz95 homo sapien
704	40	88.9	963	2	Q8BJK3_MOUSE	Q8bjk3 mus musculu	777	40	88.9	1442	2	Q7Q1Y3_ANOGA	Q7q1y3 anopheles g
705	40	88.9	963	2	Q6PGH8_MOUSE	Q6pgh8 mus musculu	778	40	88.9	1458	2	Q5LU00_MAGGR	Q5lu00 magnaporthie
706	40	88.9	968	2	Q9HCW7_HUMAN	Q9hcw7 homo sapien	779	40	88.9	1500	2	Q7XWU1_ORYSA	Q7xwu1 oryza sativ
707	40	88.9	972	2	Q51225_MAGGR	Q51225 magnaporthie	780	40	88.9	1581	2	Q6PDJ2_MOUSE	Q6pdj2 mus musculu
708	40	88.9	980	2	Q5ASA5_EMENI	Q5asa5 aspergillus	781	40	88.9	1601	2	Q5K9V7_CRYNE	Q5k9v7 cryptococcu
709	40	88.9	987	2	Q59VG5_CANAL	Q59vg5 candida alb	782	40	88.9	1604	2	Q55JP5_CRYNE	Q55jp5 cryptococcu
710	40	88.9	987	2	Q59VK1_CANAL	Q59vk1 candida alb	783	40	88.9	1609	2	Q7RY82_NEUCR	Q7ry82 neurospora
711	40	88.9	994	2	Q9V4S5_DROME	Q9v4s5 drosophila	784	40	88.9	1715	2	Q4Q7H2_LEIMA	Q4q7h2 leishmania
712	40	88.9	995	2	Q4IQ76_GIBZE	Q4iq76 gibberella	785	40	88.9	1768	2	Q7S8D6_NEUCR	Q7s8d6 neurospora
713	40	88.9	996	2	Q5CKJ5_CRYHO	Q5ckj5 cryptospori	786	40	88.9	1840	2	Q5SRY3_MOUSE	Q5sry3 mus musculu
714	40	88.9	999	2	Q4SSJ78_TETNG	Q4ss78 tetraodon n	787	40	88.9	1846	2	Q86VS8_HUMAN	Q86vs8 homo sapien
715	40	88.9	1003	2	Q4WXI4_ASPTU	Q4wx14 aspergillus	788	40	88.9	1848	2	Q69ZN8_MOUSE	Q69zn8 mus musculu
716	40	88.9	1004	2	Q4QHS6_LEIMA	Q4qhs6 leishmania	789	40	88.9	1889	1	RA11_MOUSE	Q61818 mus musculu
717	40	88.9	1023	2	Q53Q63_ORYSA	Q53q63 oryza sativ	790	40	88.9	1889	2	Q5QZ26_MOUSE	Q5qz26 mus musculu
718	40	88.9	1024	2	Q4WXM8_ASPTU	Q4wxm8 aspergillus	791	40	88.9	1889	2	Q5SRY2_MOUSE	Q5sry2 mus musculu
719	40	88.9	1036	1	STAU_DROME	P25159 drosophila	792	40	88.9	1906	1	RA11_HUMAN	Q725j4 homo sapien
720	40	88.9	1036	2	Q7SD91_NEUCR	Q7sd91 neurospora	793	40	88.9	1907	2	Q5ICW5_BRARE	Q5icw5 brachydanio
721	40	88.9	1048	2	Q5EB12_HUMAN	Q5eb12 homo sapien	794	40	88.9	1949	2	Q8UZB6_9VIRU	Q8uzb6 grapevine f
722	40	88.9	1056	2	Q6PDK1_MOUSE	Q6pdk1 mus musculu	795	40	88.9	1966	2	Q8NXX6_DROME	Q8nxx6 drosophila
723	40	88.9	1057	2	Q9EQN4_MOUSE	Q9eqn4 mus musculu	796	40	88.9	1966	2	Q8IQA6_DROME	Q8iqa6 drosophila
724	40	88.9	1058	2	Q69Z44_MOUSE	Q69z44 mus musculu	797	40	88.9	1985	2	Q9VSK5_DROSOPHILA	Q9vsk5 drosophila
725	40	88.9	1059	2	Q7S0G4_NEUCR	Q7s0g4 neurospora	798	40	88.9	1985	2	Q8T9N4_DROME	Q8t9n4 drosophila
726	40	88.9	1059	2	Q8CSX1_MOUSE	Q8csx1 mus musculu	799	40	88.9	1985	2	Q7KUAB_DROME	Q7kuab drosophila
727	40	88.9	1059	2	Q6PAL5_MOUSE	Q6pal5 mus musculu	800	40	88.9	1988	2	Q86BH2_DROME	Q86bh2 drosophila
728	40	88.9	1062	2	Q63ZY4_HUMAN	Q63zy4 homo sapien	801	40	88.9	1990	2	Q9U8Q0_DROME	Q9u8q0 drosophila
729	40	88.9	1063	2	Q6ZS14_HUMAN	Q6zsl4 homo sapien	802	40	88.9	1995	2	Q9WZ44_DROME	Q9wz44 drosophila
730	40	88.9	1070	1	Y03555_HUMAN	O15063 homo sapien	803	40	88.9	2021	2	Q4RFP3_TETNG	Q4rfp3 tetraodon n
731	40	88.9	1070	2	Q5RAJ3_PONPY	Q5raj3 pongo pygma	804	40	88.9	2053	2	Q4QDU5_LEIMA	Q4qdu5 leishmania
732	40	88.9	1070	2	Q5REB5_PONPY	Q5reb5 pongo pygma	805	40	88.9	2075	2	Q9VXY2_DROME	Q9vxy2 drosophila
733	40	88.9	1070	2	Q4R3A1_MACFA	Q4r3a1 macaca fasc	806	40	88.9	2185	2	Q75MN6_HUMAN	Q75mn6 homo sapien
734	40	88.9	1075	1	ATX2L1_HUMAN	Q8wm7 homo sapien	807	40	88.9	2220	2	Q4PBQ9_USTMA	Q4pbq9 ustilago ma
735	40	88.9	1078	2	Q59MQ9_CANAL	Q59mq9 candida alb	808	40	88.9	2246	2	Q6J514_BRARE	Q6j514 brachydanio
736	40	88.9	1078	2	Q8J2J5_CANAL	Q8j2j5 candida alb	809	40	88.9	2274	2	Q54W13_DICDI	Q54w13 dictyosteli
737	40	88.9	1085	2	Q4QHD4_LEIMA	Q4qhd4 leishmania	810	40	88.9	2343	2	Q4QIK3_LEIMA	Q4qik3 leishmania
738	40	88.9	1092	2	Q8CFT2_MOUSE	Q8cft2 mus musculu	811	40	88.9	2403	2	Q68CP0_HUMAN	Q68cp0 homo sapien
739	40	88.9	1094	2	Q5B395_EMENI	Q5b395 aspergillus	812	40	88.9	2476	2	Q4RPG5_TETNG	Q4rpg5 tetraodon n
740	40	88.9	1096	2	Q86FP1_BACOL	Q86fp1 bacrocera	813	40	88.9	2506	2	Q63HJ5_HUMAN	Q63hj5 homo sapien
741	40	88.9	1105	1	DP0D1_ORYSA	Q9lre6 oryza sativ	814	40	88.9	2563	2	Q54LJ4_DICDI	Q54lj4 dictyosteli
742	40	88.9	1105	2	Q53P50_ORYSA	Q53p50 oryza sativ	815	40	88.9	2592	2	Q9PJJ0_NEUCR	Q9pjjo neurospora
743	40	88.9	1117	2	Q8NCN1_HUMAN	Q8ncn1 homo sapien	816	40	88.9	2643	2	Q411G9_GIBZE	Q411g9 gibberella
744	40	88.9	1132	2	Q86Y26_HUMAN	Q86y26 homo sapien	817	40	88.9	2646	2	Q81220_PLAF7	Q81220 plasmodium
745	40	88.9	1134	2	Q96JH1_HUMAN	Q96jh1 homo sapien	818	40	88.9	2938	2	Q9VEE5_DROME	Q9vee5 drosophila
746	40	88.9	1146	2	Q86XA7_HUMAN	Q86xa7 homo sapien	819	40	88.9	3072	1	EP400_MOUSE	Q8chl8 mus musculu
747	40	88.9	1153	2	Q8MQW5_DROME	Q8mqw5 drosophila	820	40	88.9	3110	1	HD_RAT	P51111 rattus norv
748	40	88.9	1158	2	Q7PTX0_ANOGA	Q7ptx0 anopheles g	821	40	88.9	3119	2	Q8IHM0_PLAF7	Q8ihm0 plasmodium
749	40	88.9	1160	2	Q4WXM4_ASPTU	Q4wxm4 aspergillus	822	40	88.9	3139	2	Q9GM99_PIG	Q9gm99 sus scrofa
750	40	88.9	1162	2	Q867K1_BACOL	Q867k1 bacrocera	823	40	88.9	3146	2	Q9VUB5_DROME	Q9vub5 drosophila
751	40	88.9	1164	2	Q6CH17_YARLI	Q6ch17 yarrowia li	824	40	88.9	3246	2	Q4QGR1_LEIMA	Q4qgr1 leishmania
752	40	88.9	1170	2	Q9STI5_DROME	Q9sti5 drosophila	825	40	88.9	3396	2	Q5YL9V_MOUSE	Q5yl9v mus musculu
753	40	88.9	1174	1	KPC1_COCHE	Q42632 cochllobolu	826	40	88.9	3889	2	Q6SSE8_CHLRE	Q6sse8 chlamydomon
754	40	88.9	1181	2	Q7SEB4_NEUCR	Q7seb4 neurospora	827	39	86.7	39	2	Q9BR93_HUMAN	Q9br93 homo sapien
755	40	88.9	1185	2	Q53PZ5_ORYSA	Q53pz5 oryza sativ	828	39	86.7	45	2	Q8W701_9CAUD	Q8w701 cyanophage
756	40	88.9	1201	2	Q24240_DROME	Q24240 drosophila	829	39	86.7	62	2	Q6Z053_ORYSA	Q6z053 oryza sativ
757	40	88.9	1211	2	Q69NX2_ORYSA	Q69nx2 oryza sativ	830	39	86.7	72	2	Q4K9D4_PSBF5	Q4k9d4 pseudomonas
758	40	88.9	1212	2	Q8MKN2_DROME	Q8mkn2 drosophila	831	39	86.7	77	2	Q9NRK3_HUMAN	Q9nrk3 homo sapien
759	40	88.9	1214	1	DGKD_HUMAN	Q16760 homo sapien	832	39	86.7	82	2	Q4TPY1_9SPHN	Q4tpy1 erythrobact
760	40	88.9	1223	2	Q54W89_DICDI	Q54w89 dictyosteli	833	39	86.7	90	2	Q4VMJ0_MANSE	Q4vmj0 manduca sex
761	40	88.9	1239	2	O01505_CABEL	O01505 caenorhabdi	834	39	86.7	94	2	Q8W6Y8_9CAUD	Q8w6y8 cyanophage

835	39	86.7	95	2	Q8LBN7_ARATH	Q8Lbn7 arabidopsis	908	39	86.7	220	2	Q5T4W7_HUMAN	Q5t4w7 homo sapien
836	39	86.7	104	2	Q9FWW1_ORYSA	Q9fww1 oryza sativ	909	39	86.7	221	2	Q6Z070_ORYSA	Q6z070 oryza sativ
837	39	86.7	106	2	Q8L522_ORYSA	Q8l522 oryza sativ	910	39	86.7	224	2	Q6L4Q0_ORYSA	Q6l4q0 oryza sativ
838	39	86.7	108	2	Q91AX2_PPAPI	Q91ax2 human papil	911	39	86.7	225	2	Q6N0Q5_RHOPA	Q6n0q5 rhodopseudo
839	39	86.7	111	2	Q6YTM7_ORYSA	Q6ytm7 oryza sativ	912	39	86.7	228	2	Q6P6A3_homo sapien	Q6p6a3 homo sapien
840	39	86.7	116	2	Q5VMR4_ORYSA	Q5vmr4 oryza sativ	913	39	86.7	231	2	Q4ILQ0_GIBBEZ	Q4ilq0 gibberella
841	39	86.7	117	2	Q57Y88_9TRYP	Q57y88 trypanosoma	914	39	86.7	231	2	Q8N922_HUMAN	Q8n922 homo sapien
842	39	86.7	117	2	Q7XHN1_ORYSA	Q7xhn1 oryza sativ	915	39	86.7	232	2	Q6H694_ORYSA	Q6h694 oryza sativ
843	39	86.7	118	2	Q5NB16_ORYSA	Q5nb16 oryza sativ	916	39	86.7	233	2	Q679P4_MOUSE	Q679p4 mus musculu
844	39	86.7	123	2	Q4NW92_9DELT	Q4nw92 anaeromyxob	917	39	86.7	234	2	Q8H0A9_ORYSA	Q8h0a9 oryza sativ
845	39	86.7	126	2	Q9LPK9_ARATH	Q9lpk9 arabidopsis	918	39	86.7	235	2	Q6ZBX1_ORYSA	Q6zbx1 oryza sativ
846	39	86.7	128	2	Q9AV28_ORYSA	Q9av28 oryza sativ	919	39	86.7	236	1	UL51_PRVKA	Q85227 pseudorabie
847	39	86.7	129	2	Q6K9A5_ORYSA	Q6k9a5 oryza sativ	920	39	86.7	236	2	Q5PP96_9ALPH	Q5pp96 suid herpes
848	39	86.7	130	2	Q842K3_ECOLI	Q842k3 escherichia	921	39	86.7	237	2	Q95441_HUMAN	Q95441 homo sapien
849	39	86.7	131	2	Q99PH8_MOUSE	Q99ph8 mus musculu	922	39	86.7	241	2	Q7NDI6_GLOVI	Q7ndi6 gloeobacter
850	39	86.7	133	2	Q7SEW5_NEUCR	Q7sew5 neurospora	923	39	86.7	242	2	Q5Z4Q0_ORYSA	Q5z4q0 oryza sativ
851	39	86.7	134	2	Q4REF29_TETNG	Q4ref29 tetraodon n	924	39	86.7	244	2	Q6ETA6_ORYSA	Q6eta6 oryza sativ
852	39	86.7	136	2	Q6Z268_ORYSA	Q6z268 oryza sativ	925	39	86.7	244	2	Q4NQD2_9DELT	Q4nqd2 anaeromyxob
853	39	86.7	136	2	Q9WNX1_9DELA	Q9wnx1 human t-lym	926	39	86.7	246	2	Q6ZYK2_9VIRU	Q6zyk2 pyrobaculum
854	39	86.7	137	2	Q05284_ECOLI	Q05284 escherichia	927	39	86.7	247	2	Q5N9K0_ORYSA	Q5n9k0 oryza sativ
855	39	86.7	137	2	Q5SPH8_BRARE	Q5sph8 brachydanio	928	39	86.7	248	2	Q85622_ECOLI	Q85622 escherichia
856	39	86.7	140	2	Q6MUN0_NEUCR	Q6mun0 neurospora	929	39	86.7	256	2	Q7DB85_ECO57	Q7db85 escherichia
857	39	86.7	147	1	SMR1_MOUSE	Q61900 mus musculu	930	39	86.7	256	2	Q5YB71_MANSE	Q5y7b1 manduca sex
858	39	86.7	148	2	Q7EYB6_ORYSA	Q7eyb6 oryza sativ	931	39	86.7	259	2	Q7SSM0_NEUCR	Q7ssm0 neurospora
859	39	86.7	148	2	Q9RS96_DEIRA	Q9rs96 deinococcus	932	39	86.7	259	2	Q5SND1_ORYSA	Q5snd1 oryza sativ
860	39	86.7	148	2	Q78D20_RATRT	Q78d20 rattus ratt	933	39	86.7	260	2	Q850Z3_ORYSA	Q850z3 oryza sativ
861	39	86.7	148	2	Q64371_RAT	Q64371 rattus norv	934	39	86.7	261	2	Q5AK47_CANAL	Q5ak47 candida alb
862	39	86.7	153	2	Q653R3_ORYSA	Q653r3 oryza sativ	935	39	86.7	264	2	Q9UPA8_HUMAN	Q9upa8 homo sapien
863	39	86.7	159	2	Q5W6Y2_ORYSA	Q5w6y2 oryza sativ	936	39	86.7	267	2	Q7XXN9_ORYSA	Q7xxn9 oryza sativ
864	39	86.7	160	2	Q9AJ06_ECOLI	Q9aj06 escherichia	937	39	86.7	270	2	Q4NQ1_9DELT	Q4nqr1 anaeromyxob
865	39	86.7	161	2	Q9MON0_ARATH	Q9mon0 arabidopsis	938	39	86.7	271	2	Q4S631_TETNG	Q4s631 tetraodon n
866	39	86.7	161	2	Q81881_ARATH	Q81881 arabidopsis	939	39	86.7	275	2	Q654Z7_ORYSA	Q654z7 oryza sativ
867	39	86.7	163	2	Q7V6A1_PROMM	Q7v6a1 prochloroco	940	39	86.7	278	1	TNFL6_RAT	P36940 rattus norv
868	39	86.7	164	2	Q9AUR6_ORYSA	Q9aur6 oryza sativ	941	39	86.7	278	2	Q9BH93_DROME	Q9bh93 drosophila
869	39	86.7	166	2	Q6Z4S5_ORYSA	Q6z4s5 oryza sativ	942	39	86.7	279	1	TNFL6_MOUSE	P41047 mus musculu
870	39	86.7	170	2	Q52584_ECOLI	Q52584 escherichia	943	39	86.7	279	2	Q7TMV9_MOUSE	Q7tmv9 mus musculu
871	39	86.7	172	2	Q9DCHO_MOUSE	Q9dcho mus musculu	944	39	86.7	279	2	Q544E9_MOUSE	Q544e9 mus musculu
872	39	86.7	174	2	Q7VPW5_CHLPN	Q7vpw5 chlamydia p	945	39	86.7	280	1	TNFL6_CERTO	Q5bdl1 c tumor nec
873	39	86.7	178	2	Q7SAN4_ASHGO	Q7san4 ashbya goss	946	39	86.7	280	1	TNFL6_MACFA	P63308 m tumor nec
874	39	86.7	179	2	Q9DDY6_MOUSE	Q9ddy6 mus musculu	947	39	86.7	280	1	TNFL6_MACMU	P63307 m tumor nec
875	39	86.7	180	2	Q5JNH2_ORYSA	Q5jnh2 oryza sativ	948	39	86.7	280	1	TNFL6_MACNE	P63306 m tumor nec
876	39	86.7	181	2	Q53328_MOUSE	Q53328 mus musculu	949	39	86.7	280	2	Q861W5_FELCA	Q861w5 felis silve
877	39	86.7	182	2	Q4YR14_PLABE	Q4yr14 plasmodium	950	39	86.7	281	1	TNFL6_HUMAN	P48023 h tumor nec
878	39	86.7	182	2	Q8CEB2_MOUSE	Q8ceb2 mus musculu	951	39	86.7	281	2	Q53ZZ1_HUMAN	Q53zz1 homo sapien
879	39	86.7	185	2	Q53R94_homo sapien	Q53r94 homo sapien	952	39	86.7	282	1	TNFL6_PIG	Q9bea8 sus scrofa
880	39	86.7	188	2	Q51Q52_MAGGR	Q51q52 magnaporthe	953	39	86.7	286	2	Q9VH99_DROME	Q9vhn99 drosophila
881	39	86.7	189	2	Q67V82_ORYSA	Q67v82 oryza sativ	954	39	86.7	288	2	Q61TJ4_CAEBR	Q61tj4 caenorhabdi
882	39	86.7	190	2	Q6IL94_DROME	Q6il94 drosophila	955	39	86.7	291	2	Q51LQ6_MAGGR	Q51lq6 magnaporthe
883	39	86.7	192	2	Q51S85_MAGGR	Q51s85 magnaporthe	956	39	86.7	291	2	Q53LV0_ORYSA	Q53lv0 oryza sativ
884	39	86.7	195	2	Q8S5W0_ORYSA	Q8s5w0 oryza sativ	957	39	86.7	292	2	Q96881_DROME	Q96881 drosophila
885	39	86.7	196	2	Q8C0B8_MOUSE	Q8c0b8 mus musculu	958	39	86.7	294	2	Q5KSR0_ECOLI	Q5ksr0 escherichia
886	39	86.7	198	2	Q7XCS7_ORYSA	Q7xc57 oryza sativ	959	39	86.7	294	2	Q4TQZ7_9SPHN	Q4tqz7 erythrobaet
887	39	86.7	198	2	Q94HV7_ORYSA	Q94hv7 oryza sativ	960	39	86.7	295	2	Q94DE7_ORYSA	Q94de7 oryza sativ
888	39	86.7	198	2	Q9S2J6_STRCO	Q9s2j6 streptomyce	961	39	86.7	296	2	Q95R37_DROME	Q95r37 drosophila
889	39	86.7	198	2	Q5TY27_BRARE	Q5ty27 brachydanio	962	39	86.7	296	2	Q9VM16_DROME	Q9vm16 drosophila
890	39	86.7	199	2	Q7YJY3_STRYP	Q7y7y3 trypanosoma	963	39	86.7	302	2	Q86U16_HUMAN	Q86u16 homo sapien
891	39	86.7	199	2	Q6YU08_ORYSA	Q6yu08 oryza sativ	964	39	86.7	306	2	P93845_PEA	P93845 pisum sativ
892	39	86.7	202	2	Q82444_9DELA	Q82444 human t-lym	965	39	86.7	306	2	Q6ENK1_ORYSA	Q6enk1 oryza sativ
893	39	86.7	204	2	Q5N7R3_ORYSA	Q5n7r3 oryza sativ	966	39	86.7	308	2	Q4S667_TETNG	Q4s667 tetraodon n
894	39	86.7	204	2	Q5WME5_ECOLI	Q5wme5 escherichia	967	39	86.7	310	2	Q75K33_ORYSA	Q75k33 oryza sativ
895	39	86.7	204	2	Q9CUA2_MOUSE	Q9cu2 mus musculu	968	39	86.7	313	2	Q7VX28_BORPE	Q7vx28 bordetella
896	39	86.7	206	2	Q8GRR7_ORYSA	Q8gr7 oryza sativ	969	39	86.7	313	2	Q7WY74_BORPA	Q7wy74 bordetella
897	39	86.7	206	2	Q52152_ECOLI	Q52152 escherichia	970	39	86.7	313	2	Q7WLC4_BORBR	Q7wlc4 bordetella
898	39	86.7	207	2	Q9ROB1_ECOLI	Q9rq1 escherichia	971	39	86.7	314	1	PITX1_HUMAN	P78337 homo sapien
899	39	86.7	208	2	Q9CUQ2_MOUSE	Q9cu2 mus musculu	972	39	86.7	314	2	Q41974_MHV68	Q41974 murid herpe
900	39	86.7	211	2	Q84PY7_ORYSA	Q84py7 oryza sativ	973	39	86.7	315	1	PITX1_MOUSE	P70314 mus musculu
901	39	86.7	213	2	Q6Z0B3_ORYSA	Q6z0b3 oryza sativ	974	39	86.7	315	1	PITX1_RAT	Q99na7 rattus norv
902	39	86.7	213	2	Q9Z7M6_CHLPN	Q9z7m6 chlamydia p	975	39	86.7	316	2	Q6ZHH2_ORYSA	Q6zhh2 oryza sativ
903	39	86.7	217	2	Q8S7F2_ORYSA	Q8s7f2 oryza sativ	976	39	86.7	318	2	Q4IL55_GIBZE	Q4il55 gibberella
904	39	86.7	219	2	Q9XWF3_CAEBL	Q9xwf3 caenorhabdi	977	39	86.7	322	2	Q851R4_ORYSA	Q851r4 oryza sativ
905	39	86.7	219	2	Q9M3G8_ARATH	Q9m3g8 arabidopsis	978	39	86.7	326	2	Q599K1_BOVIN	Q599k1 bos taurus
906	39	86.7	219	2	Q994D5_PADEN	Q994d5 porcine ade	979	39	86.7	327	2	Q99JK6_MOUSE	Q99jk6 mus musculu
907	39	86.7	220	2	Q96030_homo sapien	Q96030 homo sapien	980	39	86.7	329	2	Q96S04_HUMAN	Q96s04 mus musculu

981 39 86.7 329 2 Q5ZDM9 ORYSA
 982 39 86.7 333 1 HME1_CHICK
 983 39 86.7 333 2 Q6Z4H6 ORYSA
 984 39 86.7 340 2 Q6Z303 ORYSA
 985 39 86.7 343 2 Q6IPN0 HUMAN
 986 39 86.7 345 2 Q6NUL8 HUMAN
 987 39 86.7 346 2 Q75225 HUMAN
 988 39 86.7 346 2 Q88624 MOUSE
 989 39 86.7 346 2 Q9JLP1 RAT
 990 39 86.7 346 2 Q9R226 MOUSE
 991 39 86.7 347 2 Q6K911 ORYSA
 992 39 86.7 348 2 Q7XTP3 ORYSA
 993 39 86.7 348 2 Q9MLF3 ARATH
 994 39 86.7 350 2 Q94BW2 ARATH
 995 39 86.7 351 2 Q94GL5 ORYSA
 996 39 86.7 351 2 Q7BQA4 MYCPH
 997 39 86.7 354 2 Q84QZ0 ORYSA
 998 39 86.7 354 2 Q94LH4 ORYSA
 999 39 86.7 355 2 Q9LGT9 ORYSA
 1000 39 86.7 359 2 Q50Q27_ENTHI

ALIGNMENTS

RESULT 1
 Q61P9 DROME
 ID Q61P9 DROME PRELIMINARY; PRT; 66 AA.
 AC Q61P9
 DT 05-JUL-2004 (T-EMBLrel. 27, Created)
 DT 05-JUL-2004 (T-EMBLrel. 27, Last sequence update)
 DE HDC17385.
 GN ORFNames=HDC17385;
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX PubMed=44709173; DOI=10.1186/gb-2003-5-1-r3;
 RA Hild M., Beckmann B., Haas S.A., Koch B., Solovyev V., Busold C.,
 RA Fellenberg R., Boutsos M., Vingron M., Sauer F., Hoheisel J.D.,
 RA Paro R.
 RT "An integrated gene annotation and transcriptional profiling approach
 RT towards the full gene content of the Drosophila genome."
 RL Genome Biol. 5:RESEARCH003.1-RESEARCH003.17(2003).
 CC -1- MISCELLANEOUS: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ third party annotation (TPA) entry.
 DR EMBL; BK003017; DAA03217.1; -; Genomic DNA.
 SQ SEQUENCE 66 AA; 7652 MW; D0C24FDE4E24C1D3 CRC64;
 Query Match 100.0%; Score 45; DB 2; Length 66;
 Best Local Similarity 100.0%; Pred. No. 76;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 RQPPPPP 7
 DB 49 RQPPPPP 55
 RESULT 2
 Q5VP16 ORYSA
 ID Q5VP16 ORYSA PRELIMINARY; PRT; 73 AA.
 AC Q5VP16
 DT 01-FEB-2005 (T-EMBLrel. 29, Created)
 DT 01-FEB-2005 (T-EMBLrel. 29, Last sequence update)
 DE Hypothetical protein OSUNBA0091E23.6 (Hypothetical protein
 DE B1096D03.50).
 GN Names=OSUNBA0091E23.6; Synonyms=B1096D03.50;
 OS Oryza sativa (japonica cultivar-group).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; Oryza.
 OX NCBI_TaxID=39947;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Sasaki T., Matsumoto T., Yamamoto K., Sakata K., Baba T., Katayose Y.,
 RA Wu J., Nilmura Y., Cheng Z., Nagamura Y., Antonio B.A., Kanamori H.,
 RA Hosokawa S., Masukawa M., Arikawa K., Chiden Y., Hayashi M.,
 RA Okamoto M., Ando T., Aoki H., Arita K., Hamada M., Harada C.,
 RA Hijishita S., Honda M., Ichikawa Y., Idonuma A., Iijima M., Ikeda M.,
 RA Ikono M., Itoh T., Itoh Y., Iwabuchi A., Kamiya K.,
 RA Karasawa W., Katagiri S., Kikuta A., Kobayashi N., Kono I.,
 RA Machita K., Maehara T., Mizuno H., Mizubayashi T., Mukai Y.,
 RA Nagasaki H., Nakashima M., Nakama Y., Nakamichi Y., Nakamura M.,
 RA Namiki N., Negishi M., Ohta I., Ono N., Saji S., Sakai K., Shibata M.,
 RA Shimokawa T., Shomura A., Song J., Takazaki Y., Terasawa K., Tsuji K.,
 RA Waki K., Yamagata H., Yamane H., Yoshiki S., Yoshihara R., Yukawa K.,
 RA Zhong H., Iwama H., Endo T., Ito H., Hahn J.H., Kim H.I., Sun M.Y.,
 RA Yano M., Jiang J., Gojobori T.;
 RT "The genome sequence and structure of rice chromosome 1.";
 RL Nature 420:312-316(2002).
 DR EMBL; AP003853; BAD68808.1; -; Genomic DNA.
 DR EMBL; AP003536; BAD68493.1; -; Genomic DNA.
 KW Hypothetical protein.
 SQ SEQUENCE 73 AA; 7805 MW; 5E2E35B132B8BAB4 CRC64;
 Query Match 100.0%; Score 45; DB 2; Length 73;
 Best Local Similarity 100.0%; Pred. No. 84;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 RQPPPPP 7
 DB 33 RQPPPPP 39
 RESULT 3
 Q5VPA7 ORYSA
 ID Q5VPA7 ORYSA PRELIMINARY; PRT; 119 AA.
 AC Q5VPA7
 DT 01-FEB-2005 (T-EMBLrel. 29, Created)
 DT 01-FEB-2005 (T-EMBLrel. 29, Last sequence update)
 DT 10-MAY-2005 (T-EMBLrel. 30, Last annotation update)
 DE Hypothetical protein P0498C03.3 (Hypothetical protein
 DE P0566A10.37).
 GN Name=P0498C03.3; Synonym=P0566A10.37;
 OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; Oryza.
 OX NCBI_TaxID=39947;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 6, PAC
 RT clone:P0498C03.1";
 RN Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 6, PAC
 RT clone:P0566A10.1";
 RN Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AP003724; BAD68718.1; -; Genomic DNA.
 DR EMBL; AP003630; BAD68611.1; -; Genomic DNA.
 KW Hypothetical protein.
 SQ SEQUENCE 119 AA; 12649 MW; D69766963F48B194 CRC64;
 Query Match 100.0%; Score 45; DB 2; Length 119;
 Best Local Similarity 100.0%; Pred. No. 14e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 RQPPPPP 7


```

Db          96 RPQPPPP 102
|||||
RESULT 4
Q6LKD2_9MURI PRELIMINARY; PRT; 125 AA.
AC Q6LKD2;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Synapsin I (Fragment).
GN Name=SYN1;
OS Rattus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Rattus.
OX NCBI_TaxID=10118;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Liver;
RX MEDLINE=90368667; PubMed=2118519;
RA Sauerwald A., Hoesche C., Oschwald R., Kilimann M.W.;
RT "The 5'-flanking region of the synapsin I gene. A G+C-rich, TATA- and
RT CAAT-less, phylogenetically conserved sequence with cell type-specific
RT promoter function.";
RL J. Biol. Chem. 265:14932-14937(1990).
DR EMBL; M55300; AAA42199.1; -; Genomic DNA.
DR GO; GO:0008021; C:synaptic vesicle; IEA.
DR GO; GO:0007269; P:neurotransmitter secretion; IEA.
DR InterPro; IPR001359; Synapsin.
DR Pfam; PF02078; Synapsin_N; 1.
DR PROSITE; PS00415; SYNAPSIN_1; 1.
FT NON_TER
SQ SEQUENCE 125 AA; 12193 MW; DE1B6A61659BFCA9 CRC64;

Query Match 100.0%; Score 45; DB 2; Length 125;
Best Local Similarity 100.0%; Pred. NO. 1.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
|||||
DB 27 RPQPPPP 33

RESULT 5
Q6VZB3_ORYSA PRELIMINARY; PRT; 132 AA.
AC Q6VZB3;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Hypothetical protein B1095H05.16 (Hypothetical protein
DE P020B10.40).
GN Name=B1095H05.16; Synonyms=P020B10.40;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsumoto T., Katayose Y.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 8, BAC
RT clone:B1099H05";
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsumoto T., Yamamoto K.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP005531; BAD05764.1; -; Genomic DNA.
DR EMBL; AP004656; BAD03335.1; -; Genomic DNA.
DR Gramene; Q6VZB3; -.
```

```

KW Hypothetical protein.
SQ SEQUENCE 132 AA; 14128 MW; B943F2A5417B9CD3 CRC64;

Query Match 100.0%; Score 45; DB 2; Length 132;
Best Local Similarity 100.0%; Pred. NO. 1.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
|||||
DB 47 RPQPPPP 53

RESULT 6
Q7F8E9_ORYSA PRELIMINARY; PRT; 165 AA.
AC Q7F8E9;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Hypothetical protein P0644B06.50 (Hypothetical protein
DE P0514G12.6).
GN Name=P0644B06.50; Synonyms=P0514G12.6;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 6, PAC
RT clone:P0644B06";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 6, PAC
RT clone:P0514G12.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP001129; BAA90640.1; -; Genomic DNA.
DR EMBL; AP000616; BAD67680.1; -; Genomic DNA.
DR Gramene; Q7F8E9; -;
KW Hypothetical protein.
SQ SEQUENCE 165 AA; 16785 MW; 2E7CD76B9017B159 CRC64;

Query Match 100.0%; Score 45; DB 2; Length 165;
Best Local Similarity 100.0%; Pred. NO. 1.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
|||||
DB 106 RPQPPPP 112

RESULT 7
Q654U7_ORYSA PRELIMINARY; PRT; 208 AA.
AC Q654U7;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein OJ1136_C11.22.
GN Name=OJ1136_C11.22;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 6, BAC
RT clone:OJ1136_C11.";
```


GN OrderedLocusNames=Bd0978;
OS Bdellovibrio bacteriovorus;
OC Bacteria; Proteobacteria; Deltaproteobacteria; Bdellovibrionales;
OC Bdellovibrionaceae; Bdellovibrio.
OX NCBI_TaxID=959;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=HD100 / DSM 50701 / ATCC 15356 / NCIB 9529;
RX PubMed=14752164; DOI=10.1126/science.1093027;
RA Rendulic S., Jagtap P., Rosinus A., Eppinger M., Baar C., Lanz C.,
RA Keller H., Lambert C., Evans K.J., Goemann A., Meyer F.,
RA Sockett R.E., Schuster S.C.;
RT "A predator unmasked: life cycle of Bdellovibrio bacteriovorus from a
RT genomic perspective";
RL Science 303:689-692(2004).
DR EMBL; BX842648; CAE78913.1; -; Genomic_DNA.
KW Complete proteome; Hydrolase; Signal.
FT SIGNAL 1 8 Potential.
SQ SEQUENCE 317 AA; 35558 MW; CC4BB5707945FD36 CRC64;
Query Match 100.0%; Score 45; DB 2; Length 317;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RPQPPPP 7
Db 300 RPQPPPP 306
RESULT 12
Q4T926 TETNG
ID Q4T926_TETNG PRELIMINARY; PRT; 328 AA.
AC Q4T926;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Chromosome undetermined SCAFT480, whole genome shotgun sequence.
DE (Fragment).
GN ORFNames=GSTENG00004514001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodonidae; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Jaillon O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anthouard V., Jubin C., Castellani L., Katinka M., Vacherie B.,
RA Blomont C., Skalli Z., Cactolico L., Poulain J., De Berardinis V.,
RA Cruaud C., Duprat S., Brotier P., Coutanceau J.P., Gouzy J.,
RA Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Wolff J.N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissenbach J., Roest Crolius H.;
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RT the early vertebrate proto-karyotype";
RL Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CAAE01007480; CAF90286.1; -; Genomic_DNA.
FT NON TER 1 1
SQ SEQUENCE 328 AA; 34890 MW; 4020AD04402C59BE CRC64;

Query Match 100.0%; Score 45; DB 2; Length 328;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RPQPPPP 7
Db 179 RPQPPPP 185
RESULT 13
Q9LHG8 ARATH
ID Q9LHG8 ARATH PRELIMINARY; PRT; 398 AA.
AC Q9LHG8;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE And human tumor susceptibility gene-like protein (Hypothetical protein
DE At3g12400) (Hypothetical protein T2E22.28).
GN Name=At3g12400; Synonyms=T2E22.28;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; eurosids II; Brassicales; Brassicaceae; Arabidopses.
OX NCBI_TaxID=3702;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=20363099; PubMed=10907853;
RA Kaneko T., Katoh T., Sato S., Nakamura Y., Asamizu E., Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 3. II.
RT Sequence features of the 4,251,695 bp regions covered by 90 P1, TAC
RT and BAC clones";
RL DNA Res. 7:217-221(2000).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Kaneko T., Katoh T., Sato S., Nakamura Y., Asamizu E., Tabata S.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,
RA Goldsmith A.D., Lee J.M., Quach H.L., Toriumi M., Yu G., Bowser L.,
RA Gardinetti P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
RA Miranda M., Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M.,
RA Seki M., Shinn P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,
RA Theologis A.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
RN [4]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22088475; PubMed=12093376;
RA Haas B.J., Volkovsky N., Town C.D., Troukhan M., Alexandrov N.,
RA Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;
RT "Full-length messenger RNA sequences greatly improve genome
RT annotation";
RL Genome Biol. 3:RESEARCH0029-RESEARCH0029(2002).
RN [5]
RP NUCLEOTIDE SEQUENCE.
RA Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
RA Feldmann K.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RN [6]
RP NUCLEOTIDE SEQUENCE.
RA Lin X., Kaul S., Town C.D., Benito M., Cressy T.H., Haas B.J., Wu D.,
RA Maiti R., Ronning C.M., Koo H., Fujii C.Y., Utterback T.R.,
RA Barnstead M.E., Bowman C.L., White O., Nierman W.C., Fraser C.M.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
RN [7]
RP NUCLEOTIDE SEQUENCE.
RA Yamada K., Chan M.M., Chang C.H., Dale J.M., Deng J.M., Hsuan V.W.,
RA Lee J.M., Quach H.L., Tang C., Toriumi M., Wallender E.K., Wong C.,
RA Wu H.C., Yu G., Yuan S., Bowser L., Carninci P., Chen H., Cheuk R.,
RA Havaehizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Kim C., Lam B., Lin J., Miranda M., Narusaka M., Nguyen M.,
RA Palm C.J., Sakurai T., Satou M., Seki M., Shinn P., Southwick A.,

RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 DR ENBL; AP002047; BAB03147.1; -; Genomic_DNA.
 DR ENBL; AY056283; AAL07132.1; -; mRNA.
 DR ENBL; AY087782; AAM65318.1; -; mRNA.
 DR ENBL; AC069474; AAG51025.1; -; Genomic_DNA.
 DR ENBL; AY133779; AAM91173.1; -; mRNA.
 DR HSP; Q99816; IKPQ.
 DR GO; GO:0004840; F:ubiquitin conjugating enzyme activity; IEA.
 DR GO; GO:0015031; P:protein transport; IEA.
 DR GO; GO:0006512; P:ubiquitin cycle; IEA.
 DR InterPro; IPR008883; Tsg101.
 DR InterPro; IPR000608; UBQ-conjugat_E2.
 DR Pfam; PF05743; Tsg101; 1.
 DR SMART; SM00212; UBCc; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 398 AA; 44715 MW; 659996P9784FDC20 CRC64;

Query Match 100.0%; Score 45; DB 2; Length 398;
 Best Local Similarity 100.0%; Pred. No. 4.4e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQPPPPP 7
 |||||
 DB 163 RQPPPPP 169

RESULT 14

RQ2D6_NEUCR
 ID Q9C2D6_NEUCR PRELIMINARY; PRT; 416 AA.
 AC Q9C2D6;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Hypothetical protein 966.290.
 GN Name=966.290;
 OS Neurospora crassa.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
 OX NCBI_TaxID=5141;
 [1]
 RN NUCLEOTIDE SEQUENCE.
 RP Schulte U., Aign V., Hohelsel J., Brandt P., Fartmann B., Holland R.,
 RA Nyakatura G., Mewes H.W., Mannhaupt G.;
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 [2]

RQ2D6_NEUCR
 ID Q9C2D6_NEUCR PRELIMINARY; PRT; 416 AA.
 AC Q9C2D6;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Hypothetical protein 966.290.
 GN Name=966.290;
 OS Neurospora crassa.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
 OX NCBI_TaxID=5141;
 [1]
 RN NUCLEOTIDE SEQUENCE.
 RP Schulte U., Aign V., Hohelsel J., Brandt P., Fartmann B., Holland R.,
 RA Nyakatura G., Mewes H.W., Mannhaupt G.;
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 [2]

Query Match 100.0%; Score 45; DB 2; Length 416;
 Best Local Similarity 100.0%; Pred. No. 4.6e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQPPPPP 7
 |||||

Db 349 RQPPPPP 355
 RESULT 15
 Q9VZC2_DROME
 ID Q9VZC2_DROME PRELIMINARY; PRT; 420 AA.
 AC Q9VZC2; Q88247;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
 DE CG15021-PA (RE17155P).
 GN Name=CG15021; ORFNames=CG15021;
 OS Drosophila melanogaster (fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 [1]
 RN NUCLEOTIDE SEQUENCE.
 RP MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Vandal M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fostler C., Gabrielian A.B., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-P., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou S., Zhu S., Zhu X., Smith H.O.,
 RT "The genome sequence of Drosophila melanogaster."
 RL Science 287:2185-2195(2000).
 [2]
 RN NUCLEOTIDE SEQUENCE.
 RP MEDLINE=22426065; PubMed=12537568;
 RA Celniker S.E., Wheeler D.A., Kronmiller B., Dugan S.P., Frise E., Hodgson A.,
 RA Patel S., Adams M., Champe M., Lavery T., Muzny D.M., Nelson C.R.,
 RA George R.A., Hoskins R.A., Lavery T., Muzny D.M., Nelson C.R.,
 RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
 RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
 RA Weinstein G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
 RT "Finishing a whole-genome shotgun: release 3 of the Drosophila
 RT melanogaster euchromatic genome sequence."
 RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
 [3]
 RN NUCLEOTIDE SEQUENCE.
 RP MEDLINE=22426070; PubMed=12537573;

```
RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J.W., Svirskas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celisner S.E.;
RT "The transposable elements of the Drosophila melanogaster euchromatin:
RT a genomics perspective.";
RL Genome Biol. 3:RESEARCH0084.1-RESEARCH0084.20(2002).
RN
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA Smith C.P., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celisner S.E., de Grey A.D.N.J., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.O.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
RT systematic review.";
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
RN
RP NUCLEOTIDE SEQUENCE.
RG Berkeley Drosophila Genome Project;
RA Celisner S., Carlson J., Wan K., Pfeiffer B., Frise E., George R.,
RA Hoskins R., Stapleton M., Pacleb J., Park S., Svirskas R., Smith E.,
RA Yu C., Rubin G.;
RT "Drosophila melanogaster release 4 sequence.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN
RP NUCLEOTIDE SEQUENCE.
RG FlyBase;
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
RN
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Berkley;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarini H., Krommiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nuncio J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celisner S.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE003481; AAF47902.2; -; Genomic_DNA.
DR ENBL; AY071124; AAL48746.1; -; nENA.
DR Ensembl; CG15021; Drosophila melanogaster.
DR FlyBase; FBgn0035544; CG15021.
DR GO; GO:0005199; F:structural constituent of cell wall; IEA.
DR InterPro; IPR003882; Pistil extensin.
DR InterPro; IPR004019; YLP_motif.
DR Pfam; PF02757; YLP_2.
DR PRINTS; PR01218; PSTEXTENSIN.
DR SQU SEQUENCE 420 AA; 42947 MW; 96D62FFDC9F996E1 CRC64;

Query Match 100.0%; Score 45; DB 2; Length 420;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPOPPPP 7
Db 205 RPOPPPP 211

RESULT 16
Q9P218 GVXN
ID Q9P218 GVXN PRELIMINARY; PRT; 449 AA.
AC Q9P218;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE ORF25.
GN Name=ORF25;
OS Xestia c-nigrum granulosis virus (XnGV) (Xestia c-nigrum
OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae; Granulovirus.
```

```
OX NCBI_TaxID=51677;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=99434230; PubMed=10502508; DOI=10.1006/viro.1999.9894;
RA Hayakawa T., Ko R., Okano K., Seong S.I., Goto C., Maeda S.;
RT "Sequence analysis of the Xestia c-nigrum granulovirus genome.";
RL Virology 262:277-297(1999).
DR EMBL; AF162221; AAF05139.1; -; Genomic DNA.
DR SQU SEQUENCE 449 AA; 51281 MW; 7B7F5666AECABF23 CRC64;

Query Match 100.0%; Score 45; DB 2; Length 449;
Best Local Similarity 100.0%; Pred. No. 5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPOPPPP 7
Db 332 RPOPPPP 338

RESULT 17
Q4POX1 USTMA
ID Q4POX1 USTMA PRELIMINARY; PRT; 551 AA.
AC Q4POX1;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Predicted protein.
GN ORFNames=UM06242.1;
OS Ustilago maydis 521.
OC Eukaryota; Fungi; Basidiomycota; Ustilaginomycetes;
OC Ustilaginomycetidae; Ustilaginales; Ustilaginaceae; Ustilago.
OX NCBI_TaxID=237631;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=521;
RA Birren B., Nusbaum C., Abebe A., Abouelleil A., Adekoya E.,
RA Alt-zahra M., Allen N., Allen T., An P., Anderson M., Anderson S.,
RA Arachchi H., Ambruster J., Bachanteang P., Baldwin J., Barry A.,
RA Bayul T., Blitshteyn B., Bloom T., Blye J., Boguslavskiy L.,
RA Borowsky M., Boukgalter B., Brunache A., Butler J., Calixte N.,
RA Calvo S., Camarata J., Campo K., Chang J., Cheshatsang Y., Citroen M.,
RA Collimore A., Considine T., Cook A., Cooke P., Corum B., Cuomo C.,
RA David R., Dawoe T., Degray S., Dodge S., Dooley K., Dorje P.,
RA Dorjee K., Dorris L., Duffey N., Dupes A., Elkins T., Engels R.,
RA Erickson J., Farina A., Faro S., Ferreira P., Fischer H., Gnerre S.,
RA Fitzgerald M., Foley K., Gage D., Galagan J., Gearin G., Gierre S.,
RA Gierke A., Goyette A., Graham J., Grandbois E., Gyaltsen K., Hafez N.,
RA Hagopian D., Hagos B., Hall J., Hatcher B., Heller A., Higgins H.,
RA Honan T., Horn A., Houde N., Hughes L., Hulme W., Husby E., Iliev I.,
RA Jaffe D., Jones C., Kamal M., Kamat A., Kamyseselis M., Karlsson E.,
RA Kells C., Kieu A., Kisher P., Kodira C., Kulbokas E., Labutti K.,
RA Lama D., Landers T., Leger J., Levine S., Lewis D., Lewis T.,
RA Lindblad-toh K., Liu X., Lokyitsang T., Lokyitsang Y., Lucien O.,
RA Lui A., Ma L.J., Mabbitt R., Macdonald J., Maclean C., Major J.,
RA Manning J., Marabella R., Maru K., Matthews C., Mauceli E.,
RA McCarthy M., McDonough S., McGhee T., Meldrim J., Meneus L.,
RA Mesirov J., Mihalev A., Mihova T., Mikkelson T., Mlenga V., Moru K.,
RA Mozes J., Mulraim L., Munson G., Naylor J., Neues C., Nguyen C.,
RA Nguyen N., Nguyen T., Nicol R., Nielsen C., Nizzari M., Norbu C.,
RA Norbu N., O'donnell P., Okoawo O., O'leary S., Omotosho B.,
RA O'neill K., Osman S., Parker S., Perrin D., Phunkhang P., Pignani B.,
RA Purcell S., Rachupka T., Ramasamy U., Rameau R., Ray V., Raymond C.,
RA Retta R., Richardson S., Rise C., Rodriguez J., Rogers J., Rogov P.,
RA Rutman M., Schupbach R., Seaman C., Settupalli S., Sharpe T.,
RA Sheridan J., Sherpa N., Shi J., Smirnov S., Smith C., Sougnuez C.,
RA Spencer B., Stalker J., Stange-thomann N., Stavropoulos S.,
RA Stetson K., Stone C., Stone S., Stubbs M., Talamas J., Tchuinga P.,
RA Tenzing P., Tesfaye S., Theodore J., Thoultsang Y., Topham K.,
RA Tenkay S., Teamlia T., Tsomo N., Vallee D., Vassiliev H.,
RA Venkataraman V., Vinson J., Vo A., Wade C., Wang S., Wangchuk T.,
RA Wangdi T., Whittaker C., Wilkinson J., Wu Y., Wyman D., Yadav S.,
RA Yang S., Yang X., Yeager S., Yee E., Young G., Zainoun J., Zembeck L.,
RA Zimmer A., Zody M., Lander E.;
```

```

RT "The genome sequence of Ustilago maydis.";
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAC01000237; EAK87122.1; -; Genomic DNA.
SQ SEQUENCE 551 AA; 61349 MW; 8928EC5E488FC91 CRC64;

Query Match 100.0%; Score 45; DB 2; Length 551;
Best Local Similarity 100.0%; Pred. No. 6.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQQPPPP 7
Db 198 RQQPPPP 204

RESULT 18
ID Q525T5_MAGGR PRELIMINARY; PRT; 587 AA.
AC Q525T5
DT 13-SRP-2005 (TrEMBLrel. 31, Created)
DT 13-SRP-2005 (TrEMBLrel. 31, Last sequence update)
DE Hypothetical protein.
GN ORFNames=MG06971.4;
OS Magnaporthe grisea 70-15.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetes incertae sedis; Magnaporthaceae; Magnaporthe.
OX NCBI_TaxID=242507;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=70-15;
RA Birren B., Nusbaum C., Abebe A., Abouelleil A., Adekoya E.,
RA Ait-zahra M., Allen T., Allen T., An P., Anderson M., Anderson S.,
RA Arachi H., Amburster J., Bachanteang P., Baldwin J., Barry A.,
RA Bayul T., Blitshstein B., Bloom T., Blye J., Boguslavskiy L.,
RA Borovsky M., Boukhalter B., Brunache A., Butler J., Calixte N.,
RA Calvo S., Camarata J., Campo K., Chang J., Cheshatsang Y., Citroen M.,
RA Collymore A., Considine T., Cook A., Cooke P., Corum B., Cuomo C.,
RA David R., Dawo T., Degray S., Dodge S., Dooley K., Dorje P.,
RA Dorjee K., Dorris L., Duffey N., Dupes A., Elkins T., Engels R.,
RA Erickson J., Farina A., Faro S., Ferreira P., Fischer H.,
RA Fitzgerald M., Foley K., Gage D., Galagan J., Gearin G., Gierre S.,
RA Gnikre A., Govette A., Graham J., Grandbois E., Gyaltsen K., Hafez N.,
RA Hagopian D., Hegos B., Hall J., Hatcher B., Heller A., Higgins H.,
RA Honan T., Horn A., Houde N., Hughes L., Hulme W., Husby E., Iliev I.,
RA Jaffe D., Jones C., Kamal M., Kamat A., Kamvysseis M., Karlsson E.,
RA Kells C., Kieu A., Kisaner P., Kodira C., Kulbokas E., Labutti K.,
RA Lama D., Landers T., Leger J., Levine S., Lewis D., Lewis T.,
RA Lindblad-toh K., Liu X., Lokitsang T., Lokitsang Y., Lucien O.,
RA Lui A., Ma L.J., Mabbitt R., Macdonald J., Maclean C., Major J.,
RA Manning J., Marabella R., Maru K., Matthews C., Mauceli E.,
RA McCarthy M., McDonough S., McGhee T., Meldrim J., Meneus L.,
RA Mesirov J., Mihalev A., Mihova T., Mikkelsen T., Mlenga V., Moru K.,
RA Mozes J., Mulrain L., Munson G., Naylor J., News C., Nguyen C.,
RA Nguyen N., Nguyen T., Nicol R., Nielsen C., Nizzari M., Norbu C.,
RA Norbu N., O'donnell P., Okoawo O., O'leary S., Omotosho B.,
RA O'Neill K., Osman S., Parker S., Parrin D., Phunkhang P., Pignani B.,
RA Purcell S., Rachupka T., Ramesan Y., Rameau R., Ray V., Raymond C.,
RA Retta R., Richardson S., Rise C., Rodriguez J., Rogers J., Rogov P.,
RA Rutman M., Schupbach R., Seaman C., Settupalli S., Sharpe T.,
RA Sheridan J., Sherpa N., Shi J., Smirnov S., Smith C., Sougnez C.,
RA Spencer B., Stalker J., Stange-chomann N., Stavropoulos S.,
RA Stetson K., Stone C., Stone S., Stubbs M., Talamas J., Tchuinga P.,
RA Tenzing P., Tesfaye S., Theodore J., Thoultsang Y., Topham K.,
RA Towey S., Teamlu T., Tsomo N., Vallee D., Vassiliev H.,
RA Venkataraman V., Vinson J., Vo A., Wade C., Wang S., Wangchuk T.,
RA Wangdi T., Whitaker C., Wilkinson J., Wu Y., Wyman D., Yadav S.,
RA Yang S., Yang X., Yeager S., Yee E., Young G., Zainoun J., Zembeck L.,
RA Zimmer A., Zody M., Lander E.;
RT "The genome sequence of Magnaporthe grisea.";
RT Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.

[2]
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=70-15;
RA Dean R., Mitchell T., Brown D., Brown D., Pan H., Thon M.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
[3]
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=70-15;
RA Zhu H., Blackmon B.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AACU01000620; EAA55314.1; -; Genomic DNA.
DR InterPro; IPR000232; HSF_DNA_bd.
DR InterPro; IPR002341; HSF_ETS_DNA_bd.
DR InterPro; IPR011991; Wing_hix_DNA_bd.
DR Pfam; PF00447; HSF_DNA_bind; 1.
DR ProDom; PD001788; HSF_DNA_bind; 1.
DR SMART; SM00415; HSF; 1.
KW DNA-binding; Hypothetical protein; Nuclear protein; Transcription.
SQ SEQUENCE 587 AA; 64390 MW; 11B6F08B03A3F936 CRC64;

Query Match 100.0%; Score 45; DB 2; Length 587;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQQPPPP 7
Db 409 RQQPPPP 415

RESULT 19
Q4TA45_TETNG
ID Q4TA45_TETNG PRELIMINARY; PRT; 624 AA.
AC Q4TA45;
DT 13-SRP-2005 (TrEMBLrel. 31, Created)
DT 13-SRP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SRP-2005 (TrEMBLrel. 31, Last annotation update)
DE Chromosome 10 SCA7449, whole genome shotgun sequence.
DE (fragment).
DE ORFNames=GSTENG0004438001;
GN Tetraodon nigroviridis (Green puffer).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99893;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Jaillon O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anthouard V., Jubin C., Castellani V., Katinka M., Vacherie B.,
RA Blemont C., Skalli Z., Cattolico L., Poullain J., De Berardinis V.,
RA Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J., Bosak S.,
RA Farra G., Lardier G., Chappie C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Volff J.N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissbach J., Roest Crolius H.;
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
the early vertebrate proto-karyotype.";
RL Nature 431:946-957(2004).
[2]
RN NUCLEOTIDE SEQUENCE.
RG Genoscope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CAAE01007449; CAF90237.1; -; Genomic DNA.

```

```

FT NON TER 624 624
SQ SEQUENCE 624 AA; 69882 MW; E186BFB8766BA752 CRC64;

Query Match 100.0%; Score 45; DB 2; Length 624;
Best Local Similarity 100.0%; Pred. NO. 6.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPOPPPP 7
Db 57 RPOPPPP 63

RESULT 20
Q5H9B0_HUMAN PRELIMINARY; PRT; 669 AA.
AC Q5H9B0;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Synapsin I.
GN Name=SYNI; ORFNames=RPI-230G1.2-002;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Wray P.;
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Grafham D.;
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z84466; CA142461.1; -; Genomic DNA.
DR EMBL; AL009172; CA142445.1; -; Genomic DNA.
DR EMBL; AL009172; CA142461.1; JOINED; Genomic DNA.
DR EMBL; Z84466; CA142445.1; JOINED; Genomic DNA.
DR SNR; Q5H9B0; 112-417.
DR GO; GO:0008021; C:synaptic vesicle; IEA.
DR GO; GO:0007269; P:neurotransmitter secretion; IEA.
DR InterPro; IPR001359; Synapsin.
DR Pfam; PF02750; Synapsin_C; 1.
DR Pfam; PF02078; Synapsin_N; 1.
DR PRINTS; PR01368; SYNAPSIN.
DR PROSITE; PS00415; SYNAPSIN 1; 1.
DR PROSITE; PS00416; SYNAPSIN 2; 1.
SQ SEQUENCE 669 AA; 70033 MW; 5E400115415D3E32 CRC64;

Query Match 100.0%; Score 45; DB 2; Length 669;
Best Local Similarity 100.0%; Pred. NO. 7.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPOPPPP 7
Db 27 RPOPPPP 33

RESULT 21
Q4G0U1_HUMAN PRELIMINARY; PRT; 687 AA.
AC Q4G0U1;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.

```

```

RC TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Brain;
RG NIH MGC Project;
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC036690; AAH36690.1; -; mRNA.
KW Hypothetical protein.
FT NON TER 1
SQ SEQUENCE 687 AA; 71080 MW; E5E157E801281CD CRC64;

Query Match 100.0%; Score 45; DB 2; Length 687;
Best Local Similarity 100.0%; Pred. NO. 7.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPOPPPP 7
Db 369 RPOPPPP 375

RESULT 22
SYNI_RAT
ID SYNI_RAT STANDARD; PRT; 704 AA.
AC P09951; Q9WUX7;
DT 01-MAR-1989 (Rel. 10, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Synapsin-1 (Synapsin 1).
GN Name=Syn1;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridea; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=89388265; PubMed=2506642;
RA Suedhof T.C., Czernik A.J., Rao H.-T., Takei K., Johnston P.A.,
RA Horiuchi A., Kanazir S.D., Wagner M.A., Perin M.S., de Camilli P.,
RA Greengard P.;
RT "Synapsins: mosaics of shared and individual domains in a family of
RT synaptic vesicle phosphoproteins.";
RL Science 245:1474-1480(1989).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=87133474; PubMed=3028773;
RA McCaffery C.A., Degennaro L.J.;
RT "Determination and analysis of the primary structure of the nerve
RT terminal specific phosphoprotein, synapsin I.";
RL EMBO J. 5:3167-3173(1986).

```

[3] CARBOHYDRATE-LINKAGE SITES.
RX MEDLINE=99313567; PubMed=10386995;
RA Cole R.N., Hart G.W.;
RT "Glycosylation sites flank phosphorylation sites on synapsin I: O-linked N-acetylglucosamine residues are localized within domains mediating synapsin I interactions.";
RL J. Neurochem. 73:418-428(1999).
[4]
RX CARBOHYDRATE-LINKAGE SITES.
RA MEDLINE=2236970; PubMed=12438562; DOI=10.1074/mcp.M200048-MCP200;
RWalls L., Vosseller K., Cole R.N., Cronshaw J.M., Matunis M.J., Hart G.W.;
RT "Mapping sites of O-GlcNAc modification using affinity tags for serine and threonine post-translational modifications.";
RL Mol. Cell. Proteomics 1:791-804(2002).
[5]
RX INTERACTIONS WITH NOS1 AND CAPON.
RA MEDLINE=21874142; PubMed=11867766; DOI=10.1073/pnas.261705799;
RJaffrey S.R., Benfenati F., Snowman A.M., Czernik A.J., Snyder S.H.;
RT "Neuronal nitric-oxide synthase localization mediated by a ternary complex with synapsin and CAPON.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:3199-3204(2002).
CC -1- FUNCTION: Neuronal phosphoprotein that coats synaptic vesicles, binds to the cytoskeleton, and is believed to function in the regulation of neurotransmitter release.
CC -1- SUBUNIT: Homodimer (By similarity). Interacts with CAPON. Forms a ternary complex with NOS1.
CC -1- SUBCELLULAR LOCATION: Synapse.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Names=IA;
CC IsoId=P09951-1; Sequence=Displayed;
CC Names=IB;
CC IsoId=P09951-2; Sequence=VSP_006318, VSP_006319;
CC -1- PTM: Substrate of at least four different protein kinases. It is probable that phosphorylation plays a role in the regulation of synapsin-1 in the nerve terminal (By similarity).
CC -1- SIMILARITY: Belongs to the synapsin family.
CC -1- CAUTION: Ref.2 sequence was incorrect at many places due to sequencing errors.

CC This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.

DR EMBL; M27812; AAA42145.1; -; mRNA.
DR EMBL; M27924; AAA42148.1; -; mRNA.
DR EMBL; X04655; CAA28353.1; ALT_SEQ; mRNA.
DR PIR; A25704; A25704.
DR PIR; A30411; A30411.
DR PDB; 1PK8; X-ray; A/B/C/D/E/F/G/H=2-421.
DR PDB; 1PX2; X-ray; A/B=2-421.
DR Ensembl; ENSRNOG00000010365; Rattus norvegicus.
DR RGD; 3797; Syn1.
DR GO; GO:0008021; C:synaptic vesicle; IDA.
DR GO; GO:0005515; F:protein binding; IPI.
DR GO; GO:0007269; P:neurotransmitter secretion; TAS.
DR InterPro; IPR001359; Synapsin.
DR PANTHER; PTHR10841; Synapsin; 1.
DR Pfam; PF02750; Synapsin_C; 1.
DR Pfam; PF02078; Synapsin_N; 1.
DR PRINTS; PR01368; SYNAPSIN.
DR PROSITE; PS00415; SYNAPSIN_1; 1.
DR PROSITE; PS00416; SYNAPSIN_2; 1.
DR 3D-structure; Actin-binding; Alternative splicing; Glycoprotein; Phosphorylation; Repeat; Synapse.
KW REGION 1 28 A; linker.
FT REGION 29 112 B; linker.
FT REGION 113 420 C; actin-binding and synaptic-vesicle binding.
FT

FT REGION 421 655 D; Pro-rich linker.
FT REGION 656 704 E.
FT MOD_RES 9 Phosphoserine (by CamK1 and PKA).
FT MOD_RES 566 Phosphoserine (by CamK2).
FT MOD_RES 603 Phosphoserine (by CamK2).
FT CARBOHYD 55 O-linked (GlcNAc).
FT CARBOHYD 56 O-linked (GlcNAc).
FT CARBOHYD 87 O-linked (GlcNAc).
FT CARBOHYD 96 O-linked (GlcNAc).
FT CARBOHYD 103 O-linked (GlcNAc).
FT CARBOHYD 261 O-linked (GlcNAc).
FT CARBOHYD 430 O-linked (GlcNAc).
FT CARBOHYD 516 O-linked (GlcNAc).
FT CARBOHYD 524 O-linked (GlcNAc).
FT CARBOHYD 562 O-linked (GlcNAc).
FT CARBOHYD 576 O-linked (GlcNAc).
FT VARSPLIC 660 O-linked (GlcNAc).
FT VARSPLIC 669 O-linked (GlcNAc).
FT VARSPLIC 704 Missing (in isoform IB).
FT /FTId=VSP_006319.
SQ SEQUENCE 704 AA; 73988 MW; 65799FEF7CFE18B5 CRC64;
Query Match 100.0%; Score 45; DB 1; Length 704;
Best Local Similarity 100.0%; Pred. No. 7.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RPQPPPP 7
DB 27 RPQPPPP 33
RESULT 23
Q95XQ7 CABEL
ID Q95XQ7 CABEL PRELIMINARY; PRT; 704 AA.
AC Q95XQ7;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN ORFNames=Y39G10AR.15;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RG The C. elegans sequencing consortium;
RT "Genome sequence of the nematode C. elegans: a platform for investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL; AC025716; AAK39604.1; -; Genomic DNA.
DR Ensembl; Y39G10AR.15; Caenorhabditis elegans.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 704 AA; 80412 MW; A2B949D86677EF8 CRC64;
Query Match 100.0%; Score 45; DB 2; Length 704;
Best Local Similarity 100.0%; Pred. No. 7.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RPQPPPP 7
DB 143 RPQPPPP 149
RESULT 24
SYN1 HUMAN
ID SYN1 HUMAN STANDARD; PRT; 705 AA.
AC F17600; O75825; Q5H9A9;
DT 01-AUG-1990 (Rel. 15, Created)
DT 13-SEP-2005 (Rel. 48, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)

DE Synapsin-1 (Synapsin I) (Brain protein 4.1).
GN Name=SYN1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA], AND ALTERNATIVE SPLICING.
RC TISSUE=Brain;
RX PubMed=15772651; DOI=10.1038/nature03440;
RA Ross M.T., Grafham D.V., Coffey A.J., Scherer S., McLay K., Muzny D.,
RA Platzer M., Howell G.R., Burrows C., Bird C.P., Frankish A.,
RA Lovell F.L., Howe K.L., Ashurst J.L., Fulton R.S., Sudbrak R., Wen G.,
RA Jones M.C., Hurler M.E., Andrews T.D., Fulton R.S., Searle S., Chen R.,
RA Ramser J., Whittaker A., Deadman R., Carter N.P., Hunt S.E., Chen R.,
RA Cree A., Gunaratne P., Haylak P., Hodgson A., Metzker M.L.,
RA Richards S., Scott G., Steffen D., Sodergren E., Wheeler D.A.,
RA Worley K.C., Ainscough R., Ambrose K.D., Ansari-Lari M.A., Arachya S.,
RA Ashwell R.I., Babbage A.K., Baggeley C.L., Ballabio A., Banerjee R.,
RA Barker G.E., Barlow K.F., Barrett I.P., Bates K.N., Beare D.M.,
RA Beasley H., Beasley O., Beck A., Bethel G., Blechschmidt K., Brady N.,
RA Bray-Allen S., Bridgeman A.M., Brown A.J., Brown M.J., Bonnin D.,
RA Bruford E.A., Buahy C., Burch P., Burford D., Burgess J., Burrill W.,
RA Burton J., Bye J.M., Carder C., Carrel L., Chako J., Chapman J.C.,
RA Chavez D., Chen E., Chen G., Chen Y., Chen Z., Chinault C.,
RA Ciccodicola A., Clark S.Y., Clarke G., Clee C.M., Clegg S.,
RA Clerc-Blankenburg K., Clifford K., Cobley V., Cole C.G., Conquer J.S.,
RA Corby N., Connor R.E., David R., Davies J., Davis C., Davis J.,
RA Delgado O., Deshazo D., Dhani P., Ding Y., Din H., Dodsworth S.,
RA Draper H., Dugan-Rocha S., Dunham A., Dunn M., Durbin K.J., Dutta I.,
RA Eades T., Ellwood M., Emery-Cohen A., Ezzington H., Evans K.L.,
RA Faulkner L., Francis P., Frankland J., Fraser A.E., Galkocz P.,
RA Gilbert J., Gill R., Gloeckner G., Gregory S.G., Gribble S.,
RA Griffiths C., Grocock R., Gu Y., Gwilliam R., Hamilton C., Hart E.A.,
RA Hawes A., Heath P.D., Heilmann K., Hennig S., Hernandez J.,
RA Hinzmann B., Ho S., Hoffa M., Howden P.J., Huckle E.J., Hume J.,
RA Hunt P.J., Hunt A.R., Isherwood J., Jacob L., Johnson D., Jones S.,
RA de Jong P.J., Joseph S.S., Keenan S., Kelly S., Kershaw J.K., Khan Z.,
RA Kioschis P., Klages S., Knights A.J., Kosiura A., Kovar-Smith C.,
RA Laird G.K., Langford C., Lawlor S., Leversha M., Lewis L., Liu W.,
RA Lloyd G., Lloyd D.M., Loulseged H., Loveland J.E., Lovell J.D.,
RA Lozado R., Lu J., Lyne R., Ma J., Maheshwari M., Matthews L.H.,
RA McDowell J., McLaren S., McMurray A., Meidl P., Meitinger T.,
RA Milne S., Miner G., Mistry S.L., Morgan M., Morris S., Mueller I.,
RA Mullikin J.C., Nguyen N., Nordsiek G., Nyakatura G., O'dell C.N.,
RA Okwuonu G., Palmer S., Pandian R., Parker D., Parrish J.,
RA Pasternak S., Patel D., Pearce A.V., Pearson D.M., Pelan S.E.,
RA Perez L., Porter K.M., Ramsey Y., Reichwald K., Rhodes S.,
RA Radler K.A., Schlesinger D., Schueler M.G., Sehra H.K.,
RA Shaw-Smith C., Shen H., Sheridan E.M., Showkeen R., Skuce C.D.,
RA Smith M.L., Sotherton E.C., Steingrubner H.E., Steward C.A., Storey R.,
RA Swann R.M., Swarbreck D., Tabor P.E., Taudien S., Taylor T.,
RA Teague B., Thomas K., Thorpe A., Timms K., Tracey A., Trevaion S.,
RA Tromans A.C., d'Urso M., Verdusco D., Villasana D., Waldron L.,
RA Wall M., Wang Q., Warren J., Warry G.L., Wei X., West A.,
RA Whitehead S.L., Whiteley M.N., Wilkinson J.E., Willey D.L.,
RA Williams G., Williams L., Williamson A., Williamson H., Wilming L.,
RA Woodmansey R.L., Wray P.W., Yen J., Zhang J., Zhou J., Zoghbi H.,
RA Zorilla S., Buck D., Reinhardt R., Poustka A., Rosenthal A.,
RA Lehrach H., Meindl A., Minx P.J., Hillier L.W., Willard H.F.,
RA Wilson R.K., Waterston R.H., Rice C.M., Vaudin M., Coulson A.,
RA Nelson D.L., Weinstock G., Sulston J.E., Durbin R., Hubbard T.,
RA Gibbs R.A., Beck S., Rogers J., Bentley D.R.,
RT "The DNA sequence of the human X chromosome."
RL Nature 434:325-337(2005).
[3]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA] OF 1-125.
RX MEDLINE=90368667; PubMed=2118519;
RA Sauerwald A., Hoesche C., Oschwald R., Kilmann M.W.;
RT "The 5'-flanking region of the synapsin I gene. A G+C-rich, TATA- and
RT CAAT-less, phylogenetically conserved sequence with cell type-specific
RT promoter function.";
RL J. Biol. Chem. 265:14932-14937(1990).
[4]
RN INVOLVEMENT IN EPILEPSY.
RP PubMed=14985377; DOI=10.1136/jmg.2003.013680;
RA Garcia C.C., Blair H.J., Seager M., Coulthard A., Tennant S.,
RA Buddles M., Curtis A., Goodship J.A.;
RT "Identification of a mutation in synapsin I, a synaptic vesicle
RT protein, in a family with epilepsy.";
RL J. Med. Genet. 41:183-187(2004).
CC -1- FUNCTION: Neuronal phosphoprotein that coats synaptic vesicles,
CC binds to the cytoskeleton, and is believed to function in the
CC regulation of neurotransmitter release. The complex formed with
CC NOS1 and CAPON proteins is necessary for specific nitric-oxid
CC functions at a presynaptic level.
CC -1- SUBUNIT: Homodimer. Interacts with CAPON. Forms a ternary complex
CC with NOS1 (by similarity).
CC -1- SUBCELLULAR LOCATION: Synapse.
CC Event=Alternative splicing; Named isoforms=2;
CC Name=IA;
CC IsoId=PI17600-1; Sequence=Displayed;
CC Name=IB;
CC IsoId=PI17600-2; Sequence=VSP_006316, VSP_006317;
CC -1- PPM: Substrate of at least four different protein kinases. It is
CC probable that phosphorylation plays a role in the regulation of
CC synapsin-1 in the nerve terminal.
CC -1- DISEASE: Defects in SYN1 are a cause of X-linked recessive
CC epilepsy associated with variable learning disabilities and
CC behavior disorders [MIM:300491].
CC -1- SIMILARITY: Belongs to the synapsin family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; M58378; AAC41930.1; -; Genomic_DNA.
DR EMBL; M58321; AAC41930.1; JOINED; Genomic_DNA.
DR EMBL; M58341; AAC41930.1; JOINED; Genomic_DNA.
DR EMBL; M58351; AAC41930.1; JOINED; Genomic_DNA.
DR EMBL; M58353; AAC41930.1; JOINED; Genomic_DNA.
DR EMBL; M58359; AAC41930.1; JOINED; Genomic_DNA.
DR EMBL; M58371; AAC41930.1; JOINED; Genomic_DNA.
DR EMBL; M58372; AAC41930.1; JOINED; Genomic_DNA.
DR EMBL; M58373; AAC41930.1; JOINED; Genomic_DNA.
DR EMBL; M58374; AAC41930.1; JOINED; Genomic_DNA.
DR EMBL; M58375; AAC41930.1; JOINED; Genomic_DNA.
DR EMBL; M58376; AAC41930.1; JOINED; Genomic_DNA.
DR EMBL; M58377; AAC41930.1; JOINED; Genomic_DNA.
DR EMBL; M58378; AAC41931.1; ALT SEQ; Genomic_DNA.
DR EMBL; M58379; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58381; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58382; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58383; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58384; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58385; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58386; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58387; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58388; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58389; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58390; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58391; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58392; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58393; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58394; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58395; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58396; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58397; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58398; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58399; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58400; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58401; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58402; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58403; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58404; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58405; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58406; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58407; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58408; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58409; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58410; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58411; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58412; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58413; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58414; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58415; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58416; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58417; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58418; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58419; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58420; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58421; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58422; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58423; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58424; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58425; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58426; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58427; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58428; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58429; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58430; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58431; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58432; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58433; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58434; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58435; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58436; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58437; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58438; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58439; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58440; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58441; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58442; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58443; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58444; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58445; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58446; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58447; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58448; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58449; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58450; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58451; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58452; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58453; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58454; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58455; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58456; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58457; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58458; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58459; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58460; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58461; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58462; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58463; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58464; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58465; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58466; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58467; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58468; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58469; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58470; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58471; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58472; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58473; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58474; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58475; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58476; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58477; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58478; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58479; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58480; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58481; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58482; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58483; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58484; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58485; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58486; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58487; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58488; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58489; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58490; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58491; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58492; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58493; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58494; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58495; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58496; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58497; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58498; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58499; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58500; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58501; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58502; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58503; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58504; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58505; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58506; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58507; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58508; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58509; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58510; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58511; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58512; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58513; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58514; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58515; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58516; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58517; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58518; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58519; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58520; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58521; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58522; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58523; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58524; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58525; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58526; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58527; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58528; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58529; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58530; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58531; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58532; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58533; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58534; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58535; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58536; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58537; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58538; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58539; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58540; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58541; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58542; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58543; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58544; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58545; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58546; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58547; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58548; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58549; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58550; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58551; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58552; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58553; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58554; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58555; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58556; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58557; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58558; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58559; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58560; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58561; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58562; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58563; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58564; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58565; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58566; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58567; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58568; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58569; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58570; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58571; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58572; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58573; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58574; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58575; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58576; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58577; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58578; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58579; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58580; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58581; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58582; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58583; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58584; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58585; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58586; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58587; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58588; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58589; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58590; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58591; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58592; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58593; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58594; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58595; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58596; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58597; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58598; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58599; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58600; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58601; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58602; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58603; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58604; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58605; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58606; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58607; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58608; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58609; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58610; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58611; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58612; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58613; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58614; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58615; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58616; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58617; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58618; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58619; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58620; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58621; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58622; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58623; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58624; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58625; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58626; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58627; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58628; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58629; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58630; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58631; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58632; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58633; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58634; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58635; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58636; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58637; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58638; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58639; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58640; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58641; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58642; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58643; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58644; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58645; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58646; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58647; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58648; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58649; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58650; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58651; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58652; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58653; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58654; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58655; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58656; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58657; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58658; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58659; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58660; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58661; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58662; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58663; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58664; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58665; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58666; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58667; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58668; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58669; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58670; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58671; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58672; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58673; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58674; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58675; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58676; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58677; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58678; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58679; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58680; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58681; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58682; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58683; AAC41931.1; JOINED; Genomic_DNA.
DR EMBL; M58684; AAC41931.1; JOINED; Genomic

```

DR EMBL; M55301; AAA60608.1; -; Genomic_DNA.
DR PIR; A35363; A35363.
DR HSP; P17599; 1AUX.
DR SMR; P17600; 112-417.
DR Ensembl; ENSG0000008056; Homo sapiens.
DR HGNC; HGNC:11494; SYNL.
DR MIM; 313440; -.
DR MIM; 300491; -.
DR GO; GO:0005215; F:transporter activity; TAS.
DR GO; GO:0007268; P:synaptic transmission; TAS.
DR InterPro; IPR001359; Synapsin.
DR PANTHER; PTHR10841; Synapsin; 1.
DR Pfam; PF02750; Synapsin_C; 1.
DR Pfam; PF02078; Synapsin_N; 1.
DR PRINTS; PR01368; SYNAPSIN.
DR PROSITE; PS00415; SYNAPSIN_1; 1.
DR PROSITE; PS00416; SYNAPSIN_2; 1.
KW Actin-binding; Alternative splicing; Epilepsy; Phosphorylation;
KW Repeat; Synapse.
FT REGION 1 28 A.
FT REGION 29 112 B; linker.
FT REGION 113 420 C; actin-binding and synaptic-vesicle
FT binding.
FT REGION 421 655 D; Pro-rich linker.
FT REGION 656 705 E.
FT MOD_RES 9 Phosphoserine (by CamK1 and PKA).
FT MOD_RES 568 Phosphoserine (by CamK2).
FT MOD_RES 605 Phosphoserine (by CamK2).
FT VARSPPLIC 661 669 NKQSQTNA -> KASPAQAQP (in isoform IB).
FT VARSPPLIC 670 705 Missing (in isoform IB).
FT CONFLICT 138 138 E -> G (in Ref. 1).
FT CONFLICT 631 631 R -> A (in Ref. 1).
SQ SEQUENCE 705 AA; 74111 MW; BE4CE46C942300B0 CRC64;

Query Match 100.0%; Score 45; DB 1; Length 705;
Best Local Similarity 100.0%; Pred. NO. 7.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
DB 27 RPQPPPP 33

RESULT 25
SYNL_BOVIN STANDARD; PRT; 706 AA.
AC P17599;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Synapsin-1 (Synapsin 1).
GN Name=SYNL;
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
OC Pecora; Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA] (ISOFORMS IA AND IB).
RC TISSUE=Brain;
RX MEDLINE=89388265; PubMed=2506642;
RA Suedhof T.C., Czernik A.J., Kao H.-T., Takei K., Johnston P.A.,
RA Horiuchi A., Kanatir S.D., Wagner M.A., Perin M.S., de Camilli P.,
RA Greengard P.;
RT "Synapsins: mosaics of shared and individual domains in a family of
RT synaptic vesicle phosphoproteins.";
RL Science 245:1474-1480(1989).
RN [2]
RP PHOSPHORYLATION SITE BY PROLINE-DIRECTED KINASE.
RX MEDLINE=90216728; PubMed=2108963;
RA Hall F.L., Mitchell J.P., Vulliamt P.R.;

```

```

RT "Phosphorylation of synapsin I at a novel site by proline-directed
RT protein kinase.";
RT J. Biol. Chem. 265:6944-6948(1990).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (2.15 ANGSTROMS) OF 112-417.
RX MEDLINE=98130589; PubMed=9463376; DOI=10.1093/emboj/17.4.977;
RA Esser L., Wang C.-R., Hosaka M., Smagula C.S., Suedhof T.C.,
RA Deisenhofer J.;
RT "Synapsin I is structurally similar to ATP-utilizing enzymes.";
RL EMBO J. 17:977-984(1998).
CC -!- FUNCTION: Neuronal phosphoprotein that coats synaptic vesicles,
CC binds to the cytoskeleton, and is believed to function in the
CC regulation of neurotransmitter release. The complex formed with
CC NOS1 and CAPON proteins is necessary for specific nitric-oxid
CC functions at a presynaptic level (by similarity).
CC -!- SUBUNIT: Homodimer. Interacts with CAPON. Forms a ternary complex
CC with NOS1 (by similarity).
CC -!- SUBCELLULAR LOCATION: Synapse.
CC -!- ALTERNATIVE PRODUCTS:
CC Names=IA;
CC IsoId=P17599-1; Sequence=Displayed;
CC Names=IB;
CC IsoId=P17599-2; Sequence=VSP_006314, VSP_006315;
CC -!- PTM: Substrate of at least four different protein kinases.
CC Phosphorylation may play a role in the regulation of synapsin-1 in
CC the nerve terminal.
CC -!- SIMILARITY: Belongs to the synapsin family.
CC This Swiss-Prot entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
DR EMBL; M27810; AAA30761.1; -; mRNA.
DR EMBL; M27811; AAA30762.1; -; mRNA.
DR PDB; 1AUX; X-ray; A/B=110-420.
DR PDB; 1AUX; X-ray; A/B=110-420.
DR GlycoSuiteDB; P17599; -.
DR InterPro; IPR01359; Synapsin.
DR PANTHER; PTHR10841; Synapsin; 1.
DR Pfam; PF02750; Synapsin_C; 1.
DR Pfam; PF02078; Synapsin_N; 1.
DR PRINTS; PR01368; SYNAPSIN.
DR PROSITE; PS00415; SYNAPSIN_1; 1.
DR PROSITE; PS00416; SYNAPSIN_2; 1.
KW 3D-structure; Actin-binding; Alternative splicing; Phosphorylation;
KW Repeat; Synapse.
FT REGION 1 28 A.
FT REGION 29 112 B; linker.
FT REGION 113 420 C; actin-binding and synaptic-vesicle
FT binding.
FT REGION 421 656 D; Pro-rich linker.
FT REGION 657 706 E.
FT MOD_RES 9 Phosphoserine (by CamK1 and PKA).
FT MOD_RES 551 Phosphoserine (by PDPK).
FT MOD_RES 568 Phosphoserine (by CamK2).
FT MOD_RES 605 Phosphoserine (by CamK2).
FT VARSPPLIC 662 670 NKQSQTNA -> KASPAQAQP (in isoform IB).
FT VARSPPLIC 671 706 Missing (in isoform IB).
FT VARSPPLIC 671 706 /FTId=VSP_006315.
FT STRAND 114 119
FT TURN 122 123
FT HELIX 126 130
FT TURN 131 132
FT STRAND 134 135
FT TURN 136 138
FT STRAND 139 146
FT TURN 149 150
FT STRAND 151 155
FT TURN 157 158

```

```

FT STRAND 161 165
FT TURN 172 173
FT STRAND 176 178
FT STRAND 182 185
FT STRAND 190 190
FT TURN 194 195
FT STRAND 197 197
FT STRAND 199 207
FT TURN 208 209
FT STRAND 212 213
FT HELIX 216 221
FT TURN 222 223
FT HELIX 225 239
FT TURN 241 243
FT STRAND 245 246
FT STRAND 250 252
FT HELIX 255 257
FT STRAND 266 270
FT TURN 275 276
FT STRAND 279 282
FT HELIX 285 296
FT TURN 297 299
FT STRAND 302 306
FT STRAND 310 319
FT TURN 320 321
FT STRAND 322 327
FT STRAND 345 347
FT HELIX 351 360
FT TURN 361 362
FT HELIX 363 365
FT STRAND 369 377
FT TURN 378 379
FT STRAND 382 388
FT TURN 390 391
FT TURN 398 398
FT HELIX 399 416
SQ SEQUENCE 706 AA; 89373750BF014340 CRC64;

Query Match 100.0%; Score 45; DB 1; Length 706;
Best Local Similarity 100.0%; Pred. No. 7.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
Db 27 RPQPPPP 33

RESULT 26
SYN1_MOUSE STANDARD; PRT; 706 AA.
AC O88935; Q62279; Q8QZT8;
DT 30-MAY-2000 (Rel. 39, Created)
DT 13-SEP-2005 (Rel. 48, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Synapsin-1 (Synapsin I).
GN Name=Syn1; Synonym=Syn-1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridea; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM IB).
RC STRAIN=C57BL/6; TISSUE=Pancreatic islets;
RX MEDLINE=99107854; PubMed=9890964; DOI=10.1074/jbc.274.4.2053;
RA Matsumoto K., Ebihara K., Yamamoto H., Tabuchi H., Fukunaga K.,
RA Yasunami M., Okubo H., Shichiri M., Miyamoto E.;
RT "Cloning from insulinoma cells of synapsin I associated with insulin
secretory granules.";
RL J. Biol. Chem. 274:2053-2059(1999).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 3).
RC TISSUE=Eye;

```

```

RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Stammen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Vallalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA] OF 1-125.
RX MEDLINE=94308086; PubMed=8034599;
RA Chin L.S., Li L., Greengard P.;
RT Neuron-specific expression of the synapsin II gene is directed by a
specific core promoter and upstream regulatory elements.";
RL J. Biol. Chem. 269:18507-18513(1994).
CC -1- FUNCTION: Neuronal phosphoprotein that coats synaptic vesicles,
CC binds to the cytoskeleton, and is believed to function in the
CC regulation of neurotransmitter release. Regulation of
CC neurotransmitter release. The complex formed with NOS1 and CAPON
CC proteins is necessary for specific nitric-oxid functions at a
CC presynaptic level.
CC -1- SUBUNIT: Homodimer. Interacts with CAPON. Forms a ternary complex
CC with NOS1 (By similarity).
CC -1- SUBCELLULAR LOCATION: Synapse.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Name=Ia;
CC IsoId=O88935-2; Sequence=Displayed;
CC Name=Ib;
CC IsoId=O88935-1; Sequence=VSP_015206, VSP_015207;
CC Name=3;
CC IsoId=O88935-3; Sequence=VSP_015205;
CC Note=No experimental confirmation available;
CC -1- PTM: Substrate of at least four different protein kinases. It is
CC probable that phosphorylation plays a role in the regulation of
CC synapsin-1 in the nerve terminal (By similarity).
CC -1- SIMILARITY: Belongs to the synapsin family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC ENBL; AF085809; AAD09833.1; -; mRNA.
CC DR ENBL; BC022954; AAH22954.1; -; mRNA.
CC DR ENBL; L32025; AAA79963.1; -; Genomic_DNA.
CC DR PIR; A53592; A53692.
CC DR HSP; P17599; IAU.
CC DR SMR; O88935; 112-417.
CC DR Ensembl; ENSMUSG00000037217; Mus musculus.
CC DR MGI; MGI:98460; Syn1.
CC DR GO; GO:0045202; C:synapse; IDA.
CC DR GO; GO:0030672; C:synaptic vesicle membrane; IDA.
CC DR GO; GO:0046983; F:protein dimerization activity; TAS.
CC DR InterPro; IPR001359; Synapsin.
CC DR PANTHER; PTHR10841; Synapsin; 1.
CC DR Pfam; PF02750; Synapsin_C; 1.
CC DR Pfam; PF02078; Synapsin_N; 1.
CC DR PRINTS; PR01368; SYNAPSIN.

```

```

DR PROSITE; PS00415; SYNAPSIN 1; 1.
KW Actin-binding; Alternative splicing; Phosphorylation; Repeat; Synapse.
FT REGION 1 28 A; linker.
FT REGION 29 112 C; actin-binding and synaptic-vesicle
FT REGION 113 420 binding.
FT REGION 421 657 D; Pro-rich linker.
FT MOD_RES 658 706 E.
FT MOD_RES 9 Phosphoserine (by CAMK1 and PKA) (By
FT MOD_RES 568 similarity).
FT MOD_RES 568 Phosphoserine (by CAMK2) (By similarity).
FT MOD_RES 605 Phosphoserine (by CAMK2) (By similarity).
FT VARSPLIC 573 600 Missing (in isoform 3).
FT VARSPLIC /FTid=VSP_015205.
FT VARSPLIC 662 670 NKSQLTNA -> KASPSQAP (in isoform Ib).
FT VARSPLIC /FTid=VSP_015206.
FT VARSPLIC 671 706 Missing (in isoform Ib).
FT VARSPLIC /FTid=VSP_015207.
FT CONFLICT 44 P -> L (in Ref. 3).
FT SEQUENCE 706 AA; 74097 MW; 04C940E68547372B CRC64;

Query Match 100.0%; Score 45; DB 1; Length 706;
Best Local Similarity 100.0%; Pred. No. 7.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
DB 27 RPQPPPP 33

RESULT 27
Q51L36 MAGGR PRELIMINARY; PRT; 721 AA.
ID Q51L36 MAGGR PRELIMINARY; PRT; 721 AA.
AC Q51L36;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=MG03055.4;
OS Magnaporthe grisea 70-15.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetes incertae sedis; Magnaporthaceae; Magnaportha.
OX NCBI_TaxID=242507;
RN 1;
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=70-15;
RA Birren B., Nusbaum C., Abebe A., Abouelleil A., Adekoya E.,
RA Ait-zahra M., Allen N., Allen T., An P., Anderson M., Anderson S.,
RA Arachchi H., Armbruster J., Bachantsang P., Baldwin J., Barry A.,
RA Bayul T., Blitshsteyn B., Bloom T., Blye J., Boguslavskiy L.,
RA Borowsky M., Boukhgalter B., Brunache A., Butler J., Calixte N.,
RA Calvo S., Camarata J., Campo K., Chang J., Cheshatsang Y., Citroen M.,
RA Collymore A., Considine T., Cook A., Cooke P., Corum B., Cuomo C.,
RA David R., Dawoe T., Degray N., Dodge S., Dooley K., Dorje P.,
RA Dorjee K., Dorris L., Duffey N., Dupes A., Elkins T., Engels R.,
RA Erickson J., Farina A., Faro S., Ferreira P., Fischer H.,
RA Fitzgerald W., Foley K., Gage D., Galagan J., Gearin G., Gherre S.,
RA Gnirke A., Goyette A., Graham J., Grandbois E., Gyaltzen K., Hafez N.,
RA Hagopian D., Hagos B., Hall J., Hatcher B., Heller A., Higgins H.,
RA Honan T., Horn A., Houde N., Hughes L., Hulme W., Husby E., Iliev I.,
RA Jaffe D., Jones C., Kanal M., Kamat A., Kamvyselis M., Karlsson B.,
RA Kalls C., Kieu A., Kiner P., Kodira C., Kulbokas E., Labutti K.,
RA Lama D., Landers T., Leger J., Levine S., Lewis D., Lewis T.,
RA Lindblad-toh K., Liu X., Lokyitsang T., Lokyitsang Y., Lucien O.,
RA Lui A., Ma L.J., Mabbitt R., Macdonald J., Maclean C., Major J.,
RA Manning J., Marabella R., Maru K., Matthews C., Mauceli E.,
RA McCarthy M., McDonough S., McGhee T., Meldrim J., Meneus L.,
RA Mesirov J., Mihalev A., Mihova T., Mikkelson T., Mlenga V., Moru K.,
RA Mozes J., Mulrain L., Munson G., Naylor J., News C., Nguyen C.,
RA Nguyen N., Nguyen T., Nicol R., Nielsen C., Nizzari M., Norbu C.,
RA Norbu N., O'donnell P., Okoawo O., O'leary S., Omotosho B.,
RA O'neill K., Osman S., Parker S., Perrin D., Phunkhang P., Piani B.,

```

```

RA Purcell S., Rachupka T., Ramasamy U., Rameau R., Ray V., Raymond C.,
RA Retta R., Richardson S., Rise C., Rodriguez J., Rogers J., Rogov P.,
RA Rutman M., Schupbach R., Seaman C., Settipalli S., Sharpe T.,
RA Sheridan J., Sherpa N., Shi J., Smirnov S., Smith C., Sougnez C.,
RA Spencer B., Stalker J., Stange-thomann N., Stavropoulos S.,
RA Stetson K., Stone C., Stone S., Stubbs M., Talamas J., Tchuinga P.,
RA Tenzing P., Teftaye S., Theodore J., Thoulutseang Y., Topham K.,
RA Towey S., Teamlia T., Tsomo N., Vallee D., Vassiliev H.,
RA Venkatakrman V., Vinson J., Vo A., Wade C., Wang S., Wangchuk T.,
RA Wangdi T., Whitaker C., Wilkinson J., Wu Y., Wyman D., Yadav S.,
RA Yang S., Yang X., Yeager S., Yee E., Young G., Zainoun J., Zembeck L.,
RA Zimmer A., Zody M., Lander E.;
RT "The genome sequence of Magnaporthe grisea.";
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
RN 2;
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=70-15;
RA Dean R., Mitchell T., Brown D., Pan H., Thon M.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
RN 3;
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=70-15;
RA Zhu H., Blackmon B.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AACU01001597; EAA47812.1; -; Genomic_DNA.
DR InterPro; IPR001138; Fungi_Trtscrp_N.
DR InterPro; IPR012344; Matrix_Hiv/RSV.
DR Pfam; PF00172; Zn_clus; 1.
DR SMART; SM00066; GAL4; 1.
DR PROSITE; PS00463; ZN2_CY6_FUNGAL_1; 1.
DR PROSITE; PS00448; ZN2_CY6_FUNGAL_2; 1.
KW DNA-binding; Hypothetical protein; Metal-binding; Nuclear protein;
KW Transcription; Transcription regulation; Zinc.
SQ SEQUENCE 721 AA; 78599 MW; CODECF788259DDE7 CRC64;

Query Match 100.0%; Score 45; DB 2; Length 721;
Best Local Similarity 100.0%; Pred. No. 7.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
DB 84 RPQPPPP 90

RESULT 28
Q4QF52 LEIMA PRELIMINARY; PRT; 1412 AA.
ID Q4QF52 LEIMA PRELIMINARY; PRT; 1412 AA.
AC Q4QF52;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Ubiquitin hydrolase, putative (Cysteine peptidase, clan ca, family
DE cl9, putative).
GN ORFNames=LmjF15_1300;
OS Leishmania major.
OC Eukaryota; Euzlenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN 1;
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Friedlin;
RA Peacock C.S., Murphy L., Ivens A.C., Berriman M., Blackwell J.,
RA Smith D., Collins M., Foster N., Harris D., Oliver K., O'Neill S.,
RA Saunders D., Seeger K., Warren T., Rajadream M., and Barrell B.G.;
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; CT005254; CAJ03358.1; -; Genomic_DNA.
KW Hydrolase.
SQ SEQUENCE 1412 AA; 155044 MW; 6657B5CCDD56122D CRC64;

Query Match 100.0%; Score 45; DB 2; Length 1412;
Best Local Similarity 100.0%; Pred. No. 1.5e+03;

```

```
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RPPQPPP 7
DB 832 RPPQPPP 838

RESULT 29
Q96004 HUMAN
ID Q96004_HUMAN PRELIMINARY; PRT; 1480 AA.
AC Q96004;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE KIAA1883 protein (Fragment).
GN Name=LMTK3; Synonyms=KIAA1883;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=21456161; PubMed=11572484;
RA Nagase T., Kikuno R., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XXI.
RT The complete sequences of 60 new cDNA clones from brain which code for
RT large proteins."
RL DNA Res. 8:179-187(2001).
DR ENBL; AB067470; BAB67776.1; -; mRNA.
DR HSP; AB08069; lJQH.
DR Ensembl; ENSG00000142235; Homo sapiens.
DR HGNC; HGNC:19295; LMTK3.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006468; F:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000719; Prot kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR InterPro; IPR008266; Tyr_kinase_AS.
DR PRINTS; PR00109; TYRKINASE.
DR PRODOM; PD000001; Prot kinase; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; UNKNOWN_1.
KW ATP-binding; Kinase; Nucleotide-binding; Transferase;
KW Tyrosine-protein kinase.
FT NON_TER 1
SQ SEQUENCE 1480 AA; 155952 MW; C6137E4A8B484C6F CRC64;

Query Match 100.0%; Score 45; DB 2; Length 1480;
Best Local Similarity 100.0%; Pred. No. 1.6e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RPPQPPP 7
DB 1162 RPPQPPP 1168

RESULT 30
MAST2 HUMAN
ID MAST2 HUMAN STANDARD; PRT; 1798 AA.
AC Q9P008; Q94899; Q5VT07; Q5VT08; Q7LGC4; Q8NDG1; Q96B94; Q9BYE8;
DT 10-MAY-2005 (Rel. 47, Created)
DT 10-MAY-2005 (Rel. 47, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Microtubule-associated serine/threonine-protein kinase 2
DE (EC 2.7.1.37).
GN Name=MAST2; Synonyms=KIAA0807; MAST205;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
```

```
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 1), SUBCELLULAR LOCATION, AND
RP INTERACTION WITH PC-LKC.
RC TISSUE=Brain;
RX MEDLINE=22112621; PubMed=12117771; DOI=10.1093/carcin/23.7.1139;
RA Okazaki N., Takahashi N., Kojima S., Masuho Y., Koga H.;
RT "Protocadherin LKC, a new candidate for a tumor suppressor of colon
RT and liver cancers, its association with contact inhibition of cell
RT proliferation."
RL Carcinogenesis 23:1139-1148(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA] (ISOFORM 1).
RG Human chromosome 1 international sequencing consortium;
RL Submitted (DEC-2004) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 1) AND NUCLEOTIDE
RP SEQUENCE [LARGE SCALE MRNA] OF 281-1798 (ISOFORM 2).
RC TISSUE=Lymph, and Placenta;
RX Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalek U., Smailus D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A., Smailus D.E.,
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 470-1797 (ISOFORM 1).
RC TISSUE=Brain;
RX MEDLINE=99087487; PubMed=9872452;
RA Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Miyajima N.,
RA Tanaka A., Kotani H., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XI.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro."
RL DNA Res. 5:277-286(1998).
RN [5]
RP SEQUENCE REVISION TO C-TERMINUS.
RA Ohara O., Suyama M., Nagase T., Ishikawa K.-I., Kikuno R.;
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
RN [6]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 833-1797 (ISOFORM 1).
RC TISSUE=Testis;
RG The German cDNA consortium;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
RN [7]
RP TISSUE SPECIFICITY.
RX PubMed=8902215;
RA Walden P.D., Millette C.F.;
RT "Increased activity associated with the MAST205 protein kinase complex
RT during mammalian spermiogenesis."
RL Biol. Reprod. 55:1039-1044(1996).
CC -1- FUNCTION: Appears to link the dystrophin/utrophin network with
CC microtubule filaments via the syntrophins. Phosphorylation of DMD
CC or UTRN may modulate their affinities for associated proteins.
CC Functions in a multi-protein complex in spermatid maturation.
CC Regulates lipopolyaccharide-induced IL-12 synthesis in
CC macrophages by forming a complex with TRAF6, resulting in the
```

CC Inhibition of TRAF6 NF-kappaB activation (By similarity).
CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -1- COFACTOR: Magnesium (By similarity).
CC -1- INTERACTION:
CC Q9BYE9:PC-LKC; NbExp-3; IntAct=EBI-493777, EBI-493793;
CC -1- SUBCELLULAR LOCATION: Cytoplasmic. Recruited to the sub-membranous
CC area on interaction with PC-LKC.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=Q6PQ08-1; Sequence=displayed;
CC Name=2;
CC IsoId=Q6PQ08-2; Sequence=VSP_051698, VSP_051699, VSP_051700;
CC Note=No experimental confirmation available;
CC -1- TISSUE SPECIFICITY: Abundant in the testis.
CC -1- PTM: Phosphorylated and ubiquitinated. N-terminal ubiquitination
CC leads to degradation of MAST2 by proteasome-mediated proteolysis.
CC N-terminal phosphorylation appears to be a prerequisite for
CC ubiquitination (By similarity).
CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
CC -1- SIMILARITY: Contains 1 PDZ (DHR) domain.
CC -1- CAUTION: Ref.1 sequence differs from that shown due to a
CC frameshift in position 1688.
CC -1- CAUTION: Ref.2 (CAH73245, CA116563 and CA121706) sequences differ
CC from that shown due to erroneous gene model prediction.
CC -1- CAUTION: Ref.3 (AAH15816) sequence differs from that shown due to
CC frameshifts in positions 1114 and 1387.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; AB047005; BAB40778.1; ALT_FRAME; mRNA.
DR EMBL; AL645480; CAH73244.1; -; Genomic DNA.
DR EMBL; AL358075; CAH73244.1; JOINED; Genomic DNA.
DR EMBL; AL603882; CAH73244.1; JOINED; Genomic DNA.
DR EMBL; AL603888; CAH73244.1; JOINED; Genomic DNA.
DR EMBL; AL645480; CAH73245.1; ALT_SEQ; Genomic DNA.
DR EMBL; AL358075; CAH73245.1; JOINED; Genomic DNA.
DR EMBL; AL603888; CAH73245.1; JOINED; Genomic DNA.
DR EMBL; AL603882; CA116217.1; -; Genomic DNA.
DR EMBL; AL358075; CA116217.1; JOINED; Genomic DNA.
DR EMBL; AL603888; CA116217.1; JOINED; Genomic DNA.
DR EMBL; AL645480; CA116217.1; JOINED; Genomic DNA.
DR EMBL; AL603888; CA116562.1; -; Genomic DNA.
DR EMBL; AL358075; CA116562.1; JOINED; Genomic DNA.
DR EMBL; AL603882; CA116562.1; JOINED; Genomic DNA.
DR EMBL; AL645480; CA116562.1; JOINED; Genomic DNA.
DR EMBL; AL603888; CA116563.1; ALT_SEQ; Genomic DNA.
DR EMBL; AL358075; CA116563.1; JOINED; Genomic DNA.
DR EMBL; AL645480; CA116563.1; JOINED; Genomic DNA.
DR EMBL; AL358075; CA121705.1; -; Genomic DNA.
DR EMBL; AL603882; CA121705.1; JOINED; Genomic DNA.
DR EMBL; AL603888; CA121705.1; JOINED; Genomic DNA.
DR EMBL; AL645480; CA121705.1; JOINED; Genomic DNA.
DR EMBL; AL358075; CA121706.1; ALT_SEQ; Genomic DNA.
DR EMBL; AL603888; CA121706.1; JOINED; Genomic DNA.
DR EMBL; AL645480; CA121706.1; JOINED; Genomic DNA.
DR EMBL; BC015816; AAH15816.2; ALT_INIT; mRNA.
DR EMBL; BC065499; AAH65499.1; -; mRNA.
DR EMBL; AB018350; BAA34527.2; -; mRNA.
DR EMBL; AL333319; CAD38775.1; -; mRNA.
DR HSPSP; P05132; IAPM.
DR IntAct; Q6PQ08; -.
DR Ensembl; ENSG00000086015; Homo sapiens.
DR HGNC; HGNC:19035; MAST2.
DR GO; GO:0005524; F:ATP binding; ISS.
DR GO; GO:000287; F:magnesium ion binding; ISS.
DR GO; GO:0006468; P:protein amino acid phosphorylation; ISS.
DR GO; GO:0045075; P:regulation of interleukin-12 biosynthesis; ISS.
DR GO; GO:0048515; P:spermatid differentiation; ISS.

DR InterPro; IPR001478; PDZ.
DR InterPro; IPR000961; Kinase C.
DR InterPro; IPR000719; Prot kinase.
DR InterPro; IPR008271; Ser_Thr_pkin_AS.
DR InterPro; IPR002290; Ser_Thr_pkinase.
DR Pfam; PF00595; PDZ; 1.
DR Pfam; PF00069; Kinase; 1.
DR Pfam; PF00433; Kinase C; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00228; PDZ; 1.
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PSS0106; PDZ; 1.
DR PROSITE; PSS0107; PROTEIN KINASE ATP; FALSE_NEG.
DR PROSITE; PSS0011; PROTEIN KINASE DOM; 1.
DR PROSITE; PSS0018; PROTEIN KINASE ST; 1.
KW Alternative splicing; ATP-binding; Cytoskeleton; Kinase; Magnesium;
KW Membrane; Metal-binding; Nucleotide-binding; Phosphorylation;
KW Serine/threonine-protein kinase; Transferase; Ub1 conjugation.
FT DOMAIN 512 785 PDZ.
FT NP_BIND 518 526 ATP (By similarity).
FT ACT_SITE 635 635 Proton acceptor (By similarity).
FT BINDING 541 541 ATP (By similarity).
FT VARSPLIC 327 396 Missing (in isoform 2).
FT VARSPLIC 1091 1113 Missing (in isoform 2).
FT VARSPLIC 1290 1386 Missing (in isoform 2).
FT CONFLICT 388 388 D -> E (in Ref. 1).
FT CONFLICT 659 659 I -> M (in Ref. 1, 3; AAH15816 and 4).
FT CONFLICT 1225 1225 Missing (in Ref. 3; AAH65499).
FT CONFLICT 1551 1551 D -> G (in Ref. 1 and 4).
SQ SEQUENCE 1798 AA; 196436 MW; DF5C92078A3451AF CRC64;

Query Match 100.0%; Score 45; DB 1; Length 1798;
Best Local Similarity 100.0%; Pred. No. 1.9e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPOPPPP 7
DB 9 RPOPPPP 15

RESULT 31
ID Q9ATK5_CHLRE PRELIMINARY; PRT; 2301 AA.
AC Q9ATK5;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PF6 protein.
OS Chlamydomonas reinhardtii.
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae;
OC Chlamydomonadales; Chlamydomonadaceae; Chlamydomonas.
OX NCBI_TaxID=3055;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=21gr;
RX MEDLINE=21149846; PubMed=11251084;
RA Rupp G., O'Toole E., Porter M.E.;
RT "The Chlamydomonas PF6 locus encodes a large alanine/proline-rich
RT polypeptide that is required for assembly of a central pair projection
RT and regulates flagellar motility."
RL Mol. Biol. Cell 12:739-751(2001).
DR EMBL; AF327876; AAK8270.1; -; Genomic DNA.
SQ SEQUENCE 2301 AA; 237480 MW; A296943A4CF89F9E CRC64;

Query Match 100.0%; Score 45; DB 2; Length 2301;
Best Local Similarity 100.0%; Pred. No. 2.5e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPOPPPP 7

Db 1266 RPQPPPP 1272

|||||

RESULT 32

MLL2 HUMAN STANDARD; PRT; 5262 AA.

AC O14686; O14687;

DT 10-OCT-2003 (Rel. 42, Created)

DT 10-OCT-2003 (Rel. 42, Last sequence update)

DT 13-SEP-2005 (Rel. 48, Last annotation update)

DE Myeloid/lymphoid or mixed-lineage leukemia protein 2 (ALL1-related protein).

DE protein.

GN Name=MLL2; Synonyms=ALLR;

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eumetazoa; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.

OC NCBI_TaxID=9606;

OX [1]

RN NUCLEOTIDE SEQUENCE (ISOFORMS 1; 2 AND 3).

RP MEDLINE=97388474; PubMed=9247308; DOI=10.1038/sj.onc.1201211;

RX Prasad R., Zhadanov A.B., Sedkov Y., Bullrich F., Druck T., Rallapalli R., Yano T., Alder H., Croce C.M., Huebner K., Mazo A., Canaan E.;

RA "Structure and expression pattern of human ALLR, a novel gene with strong homology to ALL-1 involved in acute leukemia and to Drosophila trithorax.";

RT Oncogene 15:549-560 (1997).

RL [2]

RN INTERACTION WITH ASC-2/NCOA6 CONTAINING COMPLEX.

RP TISSUE=Cervical carcinoma;

RC MEDLINE=22371496; PubMed=12482968; DOI=10.1128/MCB.23.1.140-149.2003;

RX Goo Y.-H., Sohn Y.C., Kim D.-H., Kim S.-W., Kang M.-J., Jung D.-J., Kwak B., Barlev N.A., Berger S.L., Chow V.T., Roeder R.G., Azorsa D.O., Meltzer P.S., Suh P.-G., Song E.J., Lee K.-J., Lee Y.C., Lee J.W.;

RA "Activating signal co-receptor 2 belongs to a novel steady-state complex that contains a subset of trithorax group proteins.";

RT Mol. Cell. Biol. 23:140-149 (2003).

RL [3]

RN PHOSPHORYLATION SITE SER-4463.

RP PubMed=15302935; DOI=10.1073/pnas.0404720101;

RX Beausoleil S.A., Jedrychowski M., Schwartz D., Elias J.E., Villen J., Li J., Cohn M.A., Cantley L.C., Gygi S.P.;

RA "Large-scale characterization of HeLa cell nuclear phosphoproteins.";

RL Proc. Natl. Acad. Sci. U.S.A. 101:12130-12135 (2004).

CC -1- FUNCTION: May be involved in transcriptional regulation.

CC -1- SUBUNIT: Belongs to the ASC-2/NCOA6 complex (ASCOM), which contains ASC-2/NCOA6, the retinoblastoma-binding protein RBQ-3/RBPS, alpha- and beta-tubulins, the trithorax group proteins MLL2 and MLL3, and ASH2/ASCL2.

CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).

CC -1- ALTERNATIVE PRODUCTS:

CC Event=Alternative splicing; Named isoforms=3;

CC Name=1;

CC IsoId=O14686-1; Sequence=Displayed;

CC Name=2;

CC IsoId=O14686-2; Sequence=VSP_008563, VSP_008559;

CC Name=3;

CC IsoId=O14686-3; Sequence=VSP_008560;

CC -1- TISSUE SPECIFICITY: Expressed in most adult tissues, including a variety of hematopoietic cells, with the exception of the liver.

CC -1- MISCELLANEOUS: This gene mapped to a chromosomal region involved in duplications and translocations associated with cancer.

CC -1- SIMILARITY: Belongs to the transcription factor trithorax family.

CC -1- SIMILARITY: Contains 5 PHD-type zinc fingers.

CC -1- SIMILARITY: Contains 1 post-SET domain.

CC -1- SIMILARITY: Contains 1 RING-type zinc finger.

CC -1- SIMILARITY: Contains 1 SET domain.

CC -----

CC This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.

CC EMBL; AF010403; AAC51734.1; -; mRNA.

CC EMBL; AF010404; AAC51735.1; -; mRNA.

CC PIR; T03454; T03454.

CC PIR; T03455; T03455.

CC HSP; O14839; IMW2.

CC Ensembl; ENSG00000167548; Homo sapiens.

CC HGNC; HGNC:7133; MLL2.

CC MIM; 602113; -.

CC GO; GO:0005634; C:nucleus; NAS.

CC GO; GO:0003700; F:transcription factor activity; TAS.

CC GO; GO:0006355; P:regulation of transcription, DNA-dependent; NAS.

CC GO; GO:0006366; P:transcription from RNA polymerase II promoter; TAS.

CC InterPro; IPR003889; FYrich C.

CC InterPro; IPR003888; FYrich N.

CC InterPro; IPR003616; PostSET.

CC InterPro; IPR001214; SET.

CC InterPro; IPR001965; Znf_PHD.

CC InterPro; IPR001841; Znf_RING.

CC Pfam; PF05965; FYR1; 1.

CC Pfam; PF05964; FYRN; 1.

CC Pfam; PF0628; PHD; 5.

CC Pfam; PF00856; SET; 1.

CC PROSITE; PS00868; POST_SET; 1.

CC PROSITE; PS0280; SET; 1.

CC PROSITE; PS01359; ZF_PHD_1; 5.

CC PROSITE; PS0016; ZF_PHD_2; 5.

CC PROSITE; PS0089; ZF_RING_2; 1.

CC Alternative splicing; Coiled coil; Metal-binding; Nuclear protein; Phosphorylation; Polymorphism; Repeat; Transcription; Transcription regulation; Zinc; Zinc-finger.

FT REPEAT 442 446 1.

FT REPEAT 460 464 2.

FT REPEAT 469 473 3.

FT REPEAT 496 500 4.

FT REPEAT 504 508 5.

FT REPEAT 521 525 6.

FT REPEAT 555 559 7.

FT REPEAT 564 568 8.

FT REPEAT 573 577 9.

FT REPEAT 582 586 10.

FT REPEAT 609 613 11.

FT REPEAT 618 622 12.

FT REPEAT 627 631 13.

FT REPEAT 645 649 14.

FT REPEAT 663 667 15.

FT DOMAIN 5121 5242 SET.

FT DOMAIN 5246 5262 Post-SET.

FT ZN_FING 226 276 PHD-type 1.

FT ZN_FING 229 274 RING-type.

FT ZN_FING 273 323 PHD-type 2.

FT ZN_FING 1102 1155 PHD-type 3.

FT ZN_FING 1152 1202 PHD-type 4.

FT ZN_FING 1229 1284 PHD-type 5.

FT REGION 439 668 15 X 5 AA repeats of S/P-P-P-E/P-E/A.

FT COILED 2397 2436 Potential.

FT COILED 2788 2809 Potential.

FT COILED 2974 3001 Potential.

FT COILED 3286 3342 Potential.

FT COILED 3437 3476 Potential.

FT COILED 3621 3701 Potential.

FT COILED 4265 4287 Potential.

FT COMPBIAS 229 326 Cys-rich.

FT COMPBIAS 374 922 Arg-rich.

FT COMPBIAS 1015 1053 Cys-rich.

FT COMPBIAS 1122 1235 Pro-rich.

FT COMPBIAS 1832 2351 Gln-rich.

FT COMPBIAS 2536 2547 Pro-rich.

FT COMPBIAS 2587 2703 Gln-rich.

FT COMPBIAS 2986 4000

```
FT COMPBIAS 3966 4085 Pro-rich.
FT COMPBIAS 4634 4702 Phosphoserine.
FT MOD RES 4463 4463 Missing (in isoform 2).
FT VARSPLIC 1 305 /FTid=VSP_008563.
FT VARSPLIC 306 672
FT PMELPAHMKACRVCRCACGAGSABLNSWFWFNSLCLC
FT HRCCHAQGGTHTSVAEQHTFVCSRFSPFPGDTPTDPDA
FT LYVACQCPKQGHVTSQPKPGPLQCEAKPLAGKAGVQLPE
FT QLEAPLNEMLPLPPPEESPLSPPEESPTSPPEASRLSP
FT PPEELPASPLPEALHLSRPLEESPLSPPEESPLSPPESS
FT PFSPLESPLSPPEESPPSALSTPLSPPEASPLSPPEPE
FT SPUSPPPELTSPPEASRLSPPEESPLSPPEESPPSP
FT PPEASRLFPPEESPLSPPEESPLSPPEASRLSPPEES
FT PMGPPPEESPMSPPEASRLSPPEESPLSPPEESPLS
FT -> MSPPPPEESPMSPPEASRLSPPEESPLSPPEESPLS
FT PPEASRLSPPEESPMSPPEESPMSPPEESPLSPPEESPLV
FT SRUSPPPEESPLSPPEESPTSPPEASRLSPPEESPTSP
FT PPDSPASPPPEESPLSPPEESPLSPPEESPLSPPEESPTSP
FT HLSPPPEESPLSPPEESPLSPPEESPLSPPEESPLSPPEESPLCAV
FT PEPSPPEESPLSPPEESPLSPPEESPLSPPEESPLSPPEESPLCAV
FT LSPQPEESPLSPPEESPLSPPEESPLSPPEESPLSPPEESPLCAV
FT EKPEEPQCPAPREELPLFPPEESPLSPPEESPLSPPEESPLCAV
FT PPLSPPEELPLSPPEESPLSPPEESPLSPPEESPLSPPEESPLCAV
FT A (in isoform 2).
FT /FTid=VSP_008559.
FT VARSPLIC 1454 1454 E -> EGCT (in isoform 3).
FT /FTid=VSP_008560.
FT VARIANT 4949 4949 R -> H (in dbSNP:3782356).
FT /FTid=VAR_017115.
FT SQ SEQUENCE 5262 AA; 564185 MW; 26B7C74CAD417E44 CRC64;
Query Match 100.0%; Score 45; DB 1; Length 5262;
Best Local Similarity 100.0%; Pred. No. 5.6e+03; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 RQPPPPP 7
Db 2111 RQPPPPP 2117
RESULT 33
Q67YU2 ARATH PRELIMINARY; PRT; 107 AA.
AC Q67YU2;
DT 25-OCT-2004 (T-EMBLrel. 28, Created)
DT 25-OCT-2004 (T-EMBLrel. 28, Last sequence update)
DE Hypothetical protein At4g21720.
GN NameAt4g21720;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Spermatophyta; Magnoliophyta; eudicotyledons; rosids;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN NUCLEOTIDE SEQUENCE.
RA Totoki Y., Seki M., Ishida J., Nakajima M., Enju A., Kamiya A.,
RA Narusaka M., Shin-i T., Nakagawa M., Sakamoto N., Oishi K., Kohara Y.,
RA Kobayashi M., Toyoda A., Sakaki Y., Sakurai T., Iida K., Akiyama K.,
RA Satou M., Toyoda T., Konagaya A., Carninci P., Kawai J.,
RA Hayashizaki Y., Shinozaki K.;
RT "Large-scale analysis of RIKEN Arabidopsis full-length (RAFL) cDNAs.";
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL035527; CAB36816.1; -; Genomic DNA.
DR EMBL; AL161555; CAB81279.1; -; Genomic DNA.
DR EMBL; BT010603; AAQ89625.1; -; mRNA.
DR EMBL; AK175223; BAD42986.1; -; mRNA.
DR PIR; T05847; T05847.
KW Hypothetical protein.
SQ SEQUENCE 107 AA; 11687 MW; 9A48D4B53F4159F4 CRC64;
Query Match 93.3%; Score 42; DB 2; Length 107;
Best Local Similarity 85.7%; Pred. No. 2.8e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 RQPPPPP 7
Db 2111 RQPPPPP 2117
RESULT 34
Q9SVS5 ARATH PRELIMINARY; PRT; 139 AA.
AC Q9SVS5;
DT 01-MAY-2000 (T-EMBLrel. 13, Created)
DT 01-MAY-2000 (T-EMBLrel. 13, Last sequence update)
DT 01-FEB-2005 (T-EMBLrel. 29, Last annotation update)
DE Hypothetical protein F17L22.180 (Hypothetical protein AT4g21720).
GN NameF17L22.180; Synonyms=At4g21720, At4g21720;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN NUCLEOTIDE SEQUENCE.
RA Bevan M., Van Der Schueren J., Chuang Y.-J., Voet M., Robben J.,
RA Volckaert G., Bancroft I., Mewes H.W., Mayer K.F.X., Schueller C.;
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
[2]
RN NUCLEOTIDE SEQUENCE.
RA EU Arabidopsis sequencing project;
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
[3]
RN NUCLEOTIDE SEQUENCE.
RA Van Der Schueren J., Vandenbussche P., Chuang Y.-J., Braeken M.,
RA Robben J., Volckaert G., Mewes H.W., Lemcke K., Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
[4]
RN NUCLEOTIDE SEQUENCE.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
[5]
RN NUCLEOTIDE SEQUENCE.
RA Cheuk R., Chen H., Kim C.-J., Shinn P., Carninci P., Hayashizaki Y.,
RA Ishida J., Kamiya A., Kawai J., Narusaka M., Sakurai T., Satou M.,
RA Seki M., Shinozaki K., Ecker J.R.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
[6]
RN NUCLEOTIDE SEQUENCE.
RA Totoki Y., Seki M., Ishida J., Nakajima M., Enju A., Kamiya A.,
RA Narusaka M., Shin-i T., Nakagawa M., Sakamoto N., Oishi K., Kohara Y.,
RA Kobayashi M., Toyoda A., Sakaki Y., Sakurai T., Iida K., Akiyama K.,
RA Satou M., Toyoda T., Konagaya A., Carninci P., Kawai J.,
RA Hayashizaki Y., Shinozaki K.;
RT "Large-scale analysis of RIKEN Arabidopsis full-length (RAFL) cDNAs.";
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL035527; CAB36816.1; -; Genomic DNA.
DR EMBL; AL161555; CAB81279.1; -; Genomic DNA.
DR EMBL; BT010603; AAQ89625.1; -; mRNA.
DR EMBL; AK175223; BAD42986.1; -; mRNA.
DR PIR; T05847; T05847.
KW Hypothetical protein.
SQ SEQUENCE 139 AA; 15436 MW; 9C01143EF81528A8 CRC64;
Query Match 93.3%; Score 42; DB 2; Length 139;
Best Local Similarity 85.7%; Pred. No. 3.6e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 RQPPPPP 7
Db 113 RQPPPPP 119
RESULT 35
Q5N912 ORYSA PRELIMINARY; PRT; 169 AA.
AC Q5N912;
DT 01-FEB-2005 (T-EMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (T-EMBLrel. 29, Last sequence update)
```



```
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Hypothetical protein P0683F02.11.
GN Name=P0683F02.11;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsumoto T., Yamamoto K., Sakata K., Baba T., Katayose Y.,
RA Wu J., Naimura Y., Cheng Z., Nagamura Y., Antonio B.A., Kanamori H.,
RA Hosokawa S., Masukawa M., Arikawa K., Chiden Y., Hayashi M.,
RA Okamoto M., Ando T., Aoki H., Arita K., Hamada M., Harada C.,
RA Hijishita S., Honda M., Ichikawa Y., Idonuma A., Iijima M., Ikeda M.,
RA Ikeno M., Itoh S., Itoh Y., Itoh Y., Iwabuchi A., Kamiya K.,
RA Kano M., Itoh S., Itoh Y., Itoh Y., Itoh Y., Iwabuchi A., Kamiya K.,
RA Kashiwagi K., Kashiwagi S., Kikuta Y., Kobayashi N., Kono I.,
RA Machita K., Maehara T., Mizuno H., Mizubayashi T., Mukai Y.,
RA Nagasaki H., Nakashima M., Nakama Y., Nakamichi Y., Nakamura M.,
RA Namiki N., Negishi M., Ohta I., Ono N., Sai S., Sakai K., Shibata M.,
RA Shimokawa T., Shomura A., Song J., Takazaki Y., Terasawa K., Teuji K.,
RA Waki K., Yamagata H., Yamane H., Yoshiki S., Yoshihara R., Yukawa K.,
RA Zhong H., Iwama H., Endo T., Ito H., Hahn J.H., Kim H.I., Eun M.Y.,
RA Yano M., Jiang J., Gojobori T.;
RT "The genome sequence and structure of rice chromosome 1.";
RL Nature 420:312-316(2002).
DR ENBL; AP003289; BAD82047.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 169 AA; 17621 MW; AC7AB908E6AE196 CRC64;

Query Match 93.3%; Score 42; DB 2; Length 169;
Best Local Similarity 85.7%; Pred. No. 4.4e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPOPPPP 7
Db 95 KPQPPPP 101

RESULT 36
Q6Z0S6 ORYSA PRELIMINARY; PRT; 246 AA.
ID Q6Z0S6 ORYSA PRELIMINARY; PRT; 246 AA.
AC Q6Z0S6
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Far-red impaired response-like protein.
GN Name=P0439B07.7; Synonyms=OJ1125.C01.31;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsumoto T., Katayose Y.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 8, PAC
clone:P0439B07.";
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 8, BAC
clone:OJ1125.C01.";
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR ENBL; AP005480; BAD10395.1; -; Genomic_DNA.
DR ENBL; AP004044; BAD33021.1; -; Genomic_DNA.
DR Gramene; Q6Z0S6; -.
SQ SEQUENCE 246 AA; 26493 MW; 7057C07BBBFFD80 CRC64;

Query Match 93.3%; Score 42; DB 2; Length 246;
Best Local Similarity 85.7%; Pred. No. 6.4e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 RPOPPPP 7
Db 203 RPEPPPP 209

RESULT 37
Q6ZQX5 HUMAN PRELIMINARY; PRT; 279 AA.
ID Q6ZQX5 HUMAN PRELIMINARY; PRT; 279 AA.
AC Q6ZQX5
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein FLJ46823.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Uterus;
RA Suzuki O., Sasaki N., Aotsuka S., Shoji T., Ichihara T., Shiohata N.,
RA Matsumoto K., Hirano M., Sano S., Momiyama H., Onogawa Y.,
RA Matsumura Y., Moriya S., Chiba E., Momiyama H., Onogawa S.,
RA Kaeriyama S., Satoh N., Matsunawa H., Takahashi E., Katsuka R.,
RA Kuga N., Kuroda A., Satoh I., Kamata K., Takami S., Terashima Y.,
RA Watanabe M., Sugiyama T., Irie R., Otsuki T., Sato H., Wakamatsu A.,
RA Ishii S., Yamamoto J., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T.,
RA Kimura K., Yamashita H., Matsuo K., Nakamura Y., Sekine M.,
RA Kikuchi H., Kanda K., Wagatsuma M., Murakawa K., Kanehori K.,
RA Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B., Suzuki Y.,
RA Sugano S., Nagahari K., Masuho Y., Nagai K., Isogai T.;
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
DR ENBL; AK128663; BAC87557.1; -; mRNA.
SQ SEQUENCE 279 AA; 30713 MW; B6200A0A0A309820 CRC64;

Query Match 93.3%; Score 42; DB 2; Length 279;
Best Local Similarity 85.7%; Pred. No. 7.2e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPOPPPP 7
Db 21 RPEPPPP 27

RESULT 38
Q96KT3 HUMAN PRELIMINARY; PRT; 287 AA.
ID Q96KT3 HUMAN PRELIMINARY; PRT; 287 AA.
AC Q96KT3
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein C8orf7 (Fragment).
GN Name=C8orf7;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Appel S., Bergheim A., Reichwald K., Reis A., Rosenthal A., Ramsay M.,
RA Hennies H.;
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR ENBL; AJ301560; CAC82736.1; -; mRNA.
DR Ensembl; ENSG00000164724; Homo sapiens.
DR HGNC; HGNC:15543; C8orf7.
KW Hypothetical protein.
FT NON TER
SQ SEQUENCE 287 AA; 30400 MW; 39234688A8B01A69 CRC64;

Query Match 93.3%; Score 42; DB 2; Length 287;
```

```
Best Local Similarity 85.7%; Pred. No. 7.4e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
DB 114 RPEPPPP 120

RESULT 39
Q7XPK1_ORYSA
ID Q7XPK1_ORYSA PRELIMINARY; PRT; 309 AA.
AC Q7XPK1;
DT 01-OCT-2003 (T-EMBLrel. 25, Created)
DT 01-OCT-2003 (T-EMBLrel. 25, Last sequence update)
DE OSJNBa0087024.9 protein.
GN Name=OSJNBa0087024.9;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoidae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]

NUCLEOTIDE SEQUENCE.
RX PubMed=12447439; DOI=10.1038/nature01183;
RA Feng Q., Zhang Y., Hao P., Wang S., Fu G., Huang Y., Li Y., Zhu J.,
RA Liu Y., Hu X., Jia P., Zhang Y., Zhao Q., Ying K., Xu S., Tang Y.,
RA Wang Q., Zhang L., Lu Y., Mu J., Lu Y., Zhang L.S., Yu Z., Fan D.,
RA Liu X., Lu T., Li C., Wu Y., Sun T., Lei H., Li T., Hu H., Guan J.,
RA Wu M., Zhang R., Zhou B., Chen Z., Chen L., Jin Z., Wang R., Yin H.,
RA Cai Z., Kang S., Lv G., Gu W., Zhu G., Tu Y., Jia J., Zhang Y.,
RA Chen J., Ren H., Chen X., Shao C., Sun Y., Hu Q., Zhang X., Zhang W.,
RA Wang L., Ding C., Sheng H., Gu J., Chen S., Ni L., Zhu F., Chen W.,
RA Lan L., Lai Y., Cheng Z., Gu M., Jiang J., Li J., Hong G., Xue Y.,
RA Han B.;
RT "Sequence and analysis of rice chromosome 4.";
RL Nature 420:316-320(2002).
DR EMBL; AL606646; CA803586.1; -; Genomic_DNA.
DR Gramene; Q7XPK1; -.
DR GO; GO:0003576; F:nucleic acid binding; IEA.
DR InterPro; IPR004087; KH.
DR InterPro; IPR004088; KH_type_1.
DR InterPro; IPR000571; ZnF_CCCH.
DR Pfam; PF001013; KH_1; 1.
DR Pfam; PF00642; zf-CCCH; 3.
DR SMART; SM00322; KH; 1.
DR SMART; SM00356; ZnF_C3H1; 3.
DR PROSITE; PS50084; KH_TYPE_1; 1.
SQ SEQUENCE 309 AA; 31788 MW; D76FP408DB12DD81 CRC64;

Query Match 93.3%; Score 42; DB 2; Length 309;
Best Local Similarity 85.7%; Pred. No. 8e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
DB 156 RPEPPPP 162

RESULT 40
Q8UDG0_AGR75
ID Q8UDG0_AGR75 PRELIMINARY; PRT; 320 AA.
AC Q8UDG0;
DT 01-JUN-2002 (T-EMBLrel. 21, Created)
DT 01-JUN-2002 (T-EMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
DE Hypothetical protein Atu2166.
GN HypothesizedName=Atu2166;
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
OX NCBI_TaxID=176299;
RN [1]
```

```
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Dupont;
RX MEDLINE=21608550; PubMed=11743193; DOI=10.1126/science.1066804;
RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,
RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
RA Kutayavin T., Levy R., Li M.-J., McClelland E., Palmieri A.,
RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Neeter E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
RT C58.";
RL Science 294:2317-2323(2001).
DR EMBL; AS009166; AAL43155.1; -; Genomic_DNA.
DR PIR; AE2842; AE2842.
DR PIR; F97619; F97619.
DR InterPro; IPR009273; DUF930.
DR Pfam; PF06059; DUF930; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 320 AA; 34447 MW; B39449CD3D1FDBD6 CRC64;

Query Match 93.3%; Score 42; DB 2; Length 320;
Best Local Similarity 85.7%; Pred. No. 8.3e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
DB 74 KPQPPPP 80

RESULT 41
Q4IEH9_GIBZE
ID Q4IEH9_GIBZE PRELIMINARY; PRT; 355 AA.
AC Q4IEH9;
DT 13-SEP-2005 (T-EMBLrel. 31, Created)
DT 13-SEP-2005 (T-EMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (T-EMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=FG04379.1;
OS Gibberella zeae PH-1.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
OX NCBI_TaxID=229533;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PH-1;
RA Birren B., Nusbaum C., Abouelleil A., Allen N., Anderson S.,
RA Arachchi H.M., Barna N., Bastien V., Bloom T., Boguslavskiy L.,
RA Boukhgalter B., Butler J., Calvo S.E., Camarata J., Chang J.,
RA Choepel Y., Collymore A., Cooke A., Cooke P., Corum B., DeArelano K.,
RA Diaz J.S., Dodge S., Dooley K., Dorris L., Elkins T., Engels R.,
RA Erickson J., Faro S., Ferreira P., FitzGerald M., Gage D., Galagan J.,
RA Gardyna S., Gnerre S., Graham L., Grand-Pierre N., Hafez N.,
RA Hagopian D., Hagos B., Hall J., Horton L., Hulme W., Iliev I.,
RA Jaffe D., Johnson R., Jones C., Kamal M., Kamat A., Karatas A.,
RA Kells C., Landers T., Levine R., Lindblad-Toh K., Liu G., Lui A.,
RA Ma L.-J., Mabbitt R., Maclean C., Macdonald P., Major J., Manning J.,
RA Matthews C., Mauceli E., McCarthy M., Meldrim J., Meneus L.,
RA Mihova T., Mlenga V., Murphy T., Naylor J., Nguyen C., Nicol R.,
RA Nielsen C.B., Norbu C., O'Connor T., O'Donnell P., O'Neill D.,
RA Oliver J., Peterson K., Phunkhang P., Pierre N., Purcell S.,
RA Rachupka A., Ramasamy U., Raymond C., Retta R., Rise C., Rogov P.,
RA Roman J., Schauer S., Schupback R., Seaman S., Severy P., Smirnov S.,
RA Smith C., Spencer B., Stange-Thomann N., Stojanovic N., Stubbs M.,
RA Talamas J., Tesfaye S., Theodore J., Topham K., Travers M.,
RA Vassiliev H., Venkataraman V.S., Viel R., Vo A., Wang S., Wilson B.,
RA Wu X., Wyman D., Young G., Zainoun J., Zembek L., Zimmer A., Zody M.,
RA Lander E.;
RT "Fusarium graminearum genome sequence.";
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
```

CC -!- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBSJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 DR EMBL; AACW01000190; EAA72760.1; -; Genomic_DNA.
 KW Hypothetical protein.
 SQ SEQUENCE 355 AA; 39535 MW; 3D967546B01D90A5 CRC64;

Query Match 93.3%; Score 42; DB 2; Length 355;
 Best Local Similarity 85.7%; Pred. No. 9.2e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPQPPPP 7
 :|||||
 Db 186 KPQPPPP 192

RESULT 42

Q7CXQ3 AGRT5
 ID Q7CXQ3 AGRT5 PRELIMINARY; PRT; 359 AA.
 AC Q7CXQ3;
 DT 05-JUN-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE AGR C 3928p.
 GN OrderedLocusNames=AGR_C 3928;
 OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
 OX NCBI_TaxID=176299;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.

RC STRAIN=Cereon;
 RA MEDLINE=21608551; PubMed=11743194; DOI=10.1126/science.1066803;
 RX Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
 RA Qurollo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
 RA Houmlel K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu P.,
 RA Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,
 RA Flanagan C., Crowell C., Gureon J., Lomo C., Sear C., Strub G.,
 RA Cielo C., Slater S.;
 RA "Genome sequence of the plant pathogen and biotechnology agent
 RT Agrobacterium tumefaciens C58.";
 RL Science 294:2323-2328 (2001).
 DR EMBL; AE008132; AAK87911.1; -; Genomic_DNA.
 DR InterPro; IPR005273; DUF930.
 DR Pfam; PF06059; DUF930; 1.
 SQ SEQUENCE 359 AA; 39061 MW; C6CF6D657C439846 CRC64;

Query Match 93.3%; Score 42; DB 2; Length 359;
 Best Local Similarity 85.7%; Pred. No. 9.3e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPQPPPP 7
 :|||||
 Db 113 KPQPPPP 119

RESULT 43

Q4SJ81 TETNG
 ID Q4SJ81 TETNG PRELIMINARY; PRT; 377 AA.
 AC Q4SJ81;
 DT 13-SEP-2005 (TrEMBLrel. 31, Created)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
 DE Chromosome 4 SCAP14575, whole genome shotgun sequence.
 DE (Fragment).
 GN ORFNames=GSTENG00017324001;
 OS Tetraodon nigroviridis (Green puffer).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
 OC Tetraodontoidea; Tetraodontidae; Tetraodon.
 OX NCBI_TaxID=99883;
 RN [1]

RP NUCLEOTIDE SEQUENCE.
 RA Jaillon O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
 RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
 RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
 RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
 RA Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
 RA Bimont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
 RA Craud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J.,
 RA Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
 RA Kellis M., Wolff J.N., Guigo R., Zody M.C., Mesirov J.,
 RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
 RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
 RA Wincker P., Lander E.S., Weissbach J., Roest Crolius H.;
 RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
 RT the early vertebrate proto-karyotype.";
 RL Nature 431:946-957 (2004).
 RN [2]

NUCLEOTIDE SEQUENCE.

RG Genoscope; Whitehead Institute Centre for Genome Research;
 RL Submitted (PEB-2004) to the EMBL/GenBank/DBJ databases.
 CC -!- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.

DR EMBL; CAAE01014575; CAF99301.1; -; Genomic_DNA.

FT NON TER 1 1

FT NON TER 377 377

SQ SEQUENCE 377 AA; 43608 MW; A272994B5E8C5770 CRC64;

Query Match 93.3%; Score 42; DB 2; Length 377;

Best Local Similarity 85.7%; Pred. No. 9.7e+02;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPQPPPP 7

||:|||||

Db 82 RPEPPPP 88

RESULT 44

Q6IC10 CAEBR
 ID Q6IC10 CAEBR PRELIMINARY; PRT; 464 AA.
 AC Q6IC10;
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE Hypothetical protein CBG13131 (Fragment).
 GN Name=CBG13131;
 OS Caenorhabditis briggsae.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6238;
 RN [1]

NUCLEOTIDE SEQUENCE.

RP The C.briggsae Sequencing Consortium;
 RG Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
 RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
 CC -!- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.

DR EMBL; CAAC01000062; CAB67589.1; -; Genomic_DNA.

DR InterPro; IPR010480; Prot_inh_I33.

DR Pfam; PF06394; Pepsin-I3; 1.

KW Hypothetical protein.

FT NON TER 464 464

SQ SEQUENCE 464 AA; 52612 MW; 986F52C4CEFE5E0F3 CRC64;

Query Match 93.3%; Score 42; DB 2; Length 464;

Best Local Similarity 85.7%; Pred. No. 1.2e+03;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPQPPPP 7

:|||||

Db 235 KPQPPPP 241

```
RESULT 45
Q9UVD1_PNECA PRELIMINARY; PRT; 493 AA.
AC Q9UVD1;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Kaxin-like serine endoprotease (Fragment).
OS Pneumocystis carinii.
OC Eukaryota; Fungi; Ascomycota; Pneumocystidomycetes; Pneumocystidaceae;
OC Pneumocystis.
OX NCBI_TaxID=4754;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=rat;
RA Russian D.A., Andrawis-Sorial V., Kovacs J.A.;
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF009222; AAF21601.1; -; mRNA.
DR HSP; P13134; 10T5.
DR MEROPS; S08.011; -.
DR GO; GO:0016021; C: integral to membrane; IEA.
DR GO; GO:0008233; F: peptidase activity; IEA.
DR GO; GO:0005199; F: structural constituent of cell wall; IEA.
DR GO; GO:0004289; F: subtilase activity; IEA.
DR GO; GO:0006508; P: proteolysis and peptidolysis; IEA.
DR InterPro; IPR002029; Pept_S8_S53.
DR InterPro; IPR003882; Pfam_1.
DR InterPro; IPR002884; Pfam_1.
DR Pfam; PF00082; Peptidase_S8; 1.
DR Pfam; PF01483; P: propeptide; 1.
DR PRINTS; PR01218; PSTEXTENSIN.
DR ProDom; PD000717; P: propeptide; 1.
DR PROSITE; PS00138; SUBTILASE_SER; 1.
KW Protease.
FT NON_TER
SQ SEQUENCE 493 AA; 54145 MW; 2D896F9D94155C0D CRC64;

Query Match 93.3%; Score 42; DB 2; Length 493;
Best Local Similarity 85.7%; Pred. No. 1.3e+03;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RQPPPPP 7
Db 401 KQPPPPP 407

RESULT 46
Q9RJX6_STRCO PRELIMINARY; PRT; 496 AA.
AC Q9RJX6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative export protein.
GN OrderedLocusNames=SC01194; ORFNames=SCG11A.25;
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomycetes.
OX NCBI_TaxID=1902;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=A3(2) / M145;
RA MEDLINE=21996410; PubMed=12000953; DOI=10.1038/417141a;
RA Bentley S.D., Chater K.F., Cardeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L.D., Oliver K., O'Neill S.,
RA Rabinowitsch E., Rajandream M.A., Rutherford K.M., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J.R., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
KW "Complete genome sequence of the model actinomycete Streptomyces

Query Match 93.3%; Score 42; DB 2; Length 493;
Best Local Similarity 85.7%; Pred. No. 1.3e+03;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RQPPPPP 7
Db 401 KQPPPPP 407

RESULT 47
Q61587_ORYSA PRELIMINARY; PRT; 542 AA.
AC Q61587;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Putative calcium-dependent protein kinase.
GN Names=OSUNBA0003C07.9;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoidae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Chow T.-Y., Hsing Y.-I.C., Chen C.-S., Chen H.-H., Liu S.-M.,
RA Chao Y.-T., Chang S.-J., Chen H.-C., Chen S.-K., Chen T.-R.,
RA Chen Y.-L., Cheng C.-H., Chung C.-I., Han S.-Y., Hsiao S.-H.,
RA Hsiung J.-N., Hsu C.-H., Huang J.-J., Kau P.-I., Lee M.-C., Leu H.-L.,
RA Li Y.-F., Lin S.-J., Lin Y.-C., Wu S.-W., Yu C.-Y., Yu S.-W.,
RA Wu H.-P., Shaw J.-F.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
CC -I- FUNCTION: May play a role in signal transduction pathways that
CC involve calcium as a second messenger (By similarity).
CC -I- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR EMBL; AC137608; AAF47064.1; -; Genomic_DNA.
DR Gramene; Q61587; -.
DR GO; GO:0005524; F: ATP binding; IEA.
DR GO; GO:0005509; F: calcium ion binding; IEA.
DR GO; GO:0004674; F: protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; F: protein-tyrosine kinase activity; IEA.
DR GO; GO:0006468; P: protein amino acid phosphorylation; IEA.
DR InterPro; IPR011992; EF-Hand type.
DR InterPro; IPR002048; EF hand Ca bd.
DR InterPro; IPR000719; Prot kinase.
DR InterPro; IPR008271; Ser Thr_pkin AS.
DR InterPro; IPR002290; Ser Thr_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00036; ehand; 4.
DR Pfam; PF00069; Pkinase; 1.
DR ProDom; PD000012; EF-hand; 2.
```

```
DR ProDom; PD000001; Prot kinase; 1.
DR SMART; SM00054; Efb; 4.
DR SMART; SM02220; S_TKc; 1.
DR SMART; SM00219; TykK; 1.
DR PROSITE; PS00018; EF_HAND; 4.
DR PROSITE; PS00107; PROTEIN KINASE ATP; UNKNOWN_1.
DR PROSITE; PS00011; PROTEIN KINASE DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR ATP-binding; Calcium; Kinase; Nucleotide-binding; Repeat;
KW Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 542 AA; 61374 MW; 5E079B20A858B519 CRC64;

Query Match 93.3%; Score 42; DB 2; Length 542;
Best Local Similarity 85.7%; Pred. No. 1.4e+03;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
Db 41 KPQPPPP 47

RESULT 48
ID Q9FMF8 ARATH PRELIMINARY; PRT; 549 AA.
AC Q9FMF8
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE 2-oxoglutarate/malate translocator.
GN NamesAt5g64280;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=98162728; PubMed=9501997;
RA Nakamura Y., Sato S., Kaneko T., Kotani H., Asamizu E., Miyajima N.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. III.
RT Sequence features of the regions of 1,191,918 bp covered by seventeen
RT physically assigned P1 clones.";
RL DNA Res. 4:401-414(1997).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,
RA Goldsmith A.D., Lee J.M., Quach H.L., Toriumi M., Yu G., Bowser L.,
RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
RA Miranda M., Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M.,
RA Seki M., Shinn P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,
RA Theologis A.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RA Yamada K., Chan M.M., Chang C.H., Dale J.M., Deng J.M., Hsuan V.W.,
RA Lee J.M., Quach H.L., Tang C., Toriumi M., Wallender E.K., Wong C.,
RA Wu H.C., Yu G., Yuan S., Bowser L., Carninci P., Chen H., Cheuk R.,
RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Kim C., Lam B., Lin J., Miranda M., Narusaka M., Nguyen M.,
RA Palm C.J., Sakurai T., Satou M., Seki M., Shinn P., Southwick A.,
RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB008268; BAB09859.1; -; Genomic_DNA.
DR EMBL; AY045893; AAK76567.1; -; mRNA.
DR EMBL; AY133825; AAM91759.1; -; mRNA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006814; P:sodium ion transport; IEA.
DR InterPro; IPR001896; Na_sulph_symp; 1.
DR Pfam; PF00939; Na_sulph_symp; 1.
DR TIGRFAMs; TIGR00785; dss; 1.

SQ SEQUENCE 549 AA; 58747 MW; 61FD036CDB0D45F4 CRC64;

Query Match 93.3%; Score 42; DB 2; Length 549;
Best Local Similarity 85.7%; Pred. No. 1.4e+03;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
Db 66 KPQPPPP 72

RESULT 49
ID Q4RP94 TETNG PRELIMINARY; PRT; 579 AA.
AC Q4RP94
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Chromosome 1 SCAF15008, whole genome shotgun sequence.
DE (Fragment).
GN CRPNames-GSTENG00031214001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Jaillon O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA Blomont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
RA Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J.,
RA Parra G., Lardier G., Chapelle C., McKernan K.J., McSwan P., Bosak S.,
RA Kellis M., Volff J.N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudat V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissbach J., Roest Crollius H.;
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RT the early vertebrate proto-karyotype.";
RL Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CRAE01015008; CAG09788.1; -; Genomic_DNA.
FT NON_TER 1
FT NON_TER 579
FT SEQUENCE 579 AA; 65337 MW; E861791CB94FE00A CRC64;

Query Match 93.3%; Score 42; DB 2; Length 579;
Best Local Similarity 85.7%; Pred. No. 1.5e+03;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
Db 140 RPSPPPPP 146

RESULT 50
ID Q6AHW2 PNECA PRELIMINARY; PRT; 619 AA.
AC Q6AHW2
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Protease-1 (Prt1), putative (fragment)).
```

GN Name=PC11A11.01;
 OS Pneumocystis carinii.
 OC Eukaryota; Fungi; Ascomycota; Pneumocystidomycetes; Pneumocystidaceae;
 OC Pneumocystis.
 OX NCBI_TaxID=4754;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX PubMed=15496914; DOI=10.1038/nature03025;
 RA Jaillon O., Aury J.-M., Brunet F., Petit J.-L., Stange-Thomann N.,
 RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
 RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
 RA Desilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
 RA Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
 RA Blumet C., Skalli Z., Cattolico L., Poulain J., Berardinis Vd.,
 RA Craud C., Duprat S., Brottier P., Coutanceau J.-P., Gouzy J.,
 RA Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
 RA Kallis M., Wolff J.-N., Guigs R., Zody M.C., Mesirov J.,
 RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
 RA Laudet V., Schachter V., Quittier F., Saurin W., Scarpelli C.,
 RA Wincker P., Lander E.S., Weissenbach J., Crolius H.R.;
 RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
 the early vertebrate proto-karyotype.";
 RL Nature 431:946-957(2004).
 RN [2]

RP NUCLEOTIDE SEQUENCE.
 RA Renauld H., Keely S.P., Barrell B., Stringer J.R., Berriman M.,
 RA Aslett M.A., Hall N.;
 RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; CR16157; CA117853.1; -; Genomic DNA.
 DR GO; GO:0008233; F:peptidase activity; IEA.
 DR GO; GO:0005199; F:structural constituent of cell wall; IEA.
 DR GO; GO:0004289; F:subtilase activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR002029; Pept_S8_S53.
 DR InterPro; IPR003882; FstII_extensin.
 DR InterPro; IPR002884; PrptnconvertP.
 DR Pfam; PF01483; P_proprotein; 1.
 DR Pfam; PF00082; Peptidase S8; 1.
 DR PRINTS; PR01218; PSTLEXTENSIN.
 DR ProDom; PD000717; PrptnconvertP; 1.
 KW Protease.
 FT NON TER
 SQ SEQUENCE 619 AA; 66355 MW; 11601A87528508D3 CRC64;

Query Match 93.3%; Score 42; DB 2; Length 619;
 Best Local Similarity 85.7%; Pred.No. 1.6e+03;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
 :|||||
 Db 528 KPQPPPP 534

Search completed: April 6, 2006, 09:39:50
 Job time : 109.105 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 6, 2006, 09:22:55 ; Search time 134.211 Seconds
(without alignments)
39.286 Million cell updates/sec

Title: US-10-632-388-307

Perfect score: 48

Sequence: 1 XXXRPLPLXP 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Database :

A Geneseq 21:*

- 1: geneseq1980s:*
- 2: geneseq1990s:*
- 3: geneseq2000s:*
- 4: geneseq2001s:*
- 5: geneseq2002s:*
- 6: geneseq2003as:*
- 7: geneseq2003bs:*
- 8: geneseq2004s:*
- 9: geneseq2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	48	100.0	12	2 AAR93346	Aar93346 SRC/LYN p
2	48	100.0	12	2 AAR93356	Aar93356 FYN/LYN p
3	48	100.0	15	2 AAR93471	Aar93471 GST-SRC p
4	47	97.9	11	3 AAB21130	Aab21130 Src homol
5	47	97.9	11	3 AAB21125	Aab21125 Src homol
6	47	97.9	12	2 AAR93351	Aar93351 FYN prote
7	47	97.9	12	2 AAR93359	Aar93359 LYN prote
8	47	97.9	12	2 AAR93344	Aar93344 SRC prote
9	47	97.9	12	3 AAB17254	Aab17254 SH3 antag
10	47	97.9	12	3 AAB17251	Aab17251 SH3 antag
11	47	97.9	12	5 ABB73244	Abb73244 Src homol
12	47	97.9	12	5 ABB73247	Abb73247 Src homol
13	47	97.9	12	7 ADJ73401	Adj73401 SH3 antag
14	47	97.9	12	7 ADJ73398	Adj73398 SH3 antag
15	47	97.9	12	8 ADJ53032	Adj53032 CH1 delet
16	47	97.9	12	8 ADJ53035	Adj53035 CH1 delet
17	47	97.9	12	8 ADJ51993	Adj51993 CH1 delet
18	47	97.9	12	8 ADJ51996	Adj51996 CH1 delet
19	47	97.9	15	2 AAR93461	Aar93461 GST-P13K
20	47	97.9	15	2 AAR93457	Aar93457 GST-LYN p
21	47	97.9	15	2 AAR93460	Aar93460 GST-LYN p
22	47	97.9	15	2 AAR93458	Aar93458 GST-LYN p
23	47	97.9	15	2 AAR93459	Aar93459 GST-LYN p
24	47	97.9	15	2 AAW05451	Aaw05451 SH3-bind1

25	47	97.9	15	2 AAW37659	Aaw37659 PPPPY mot
26	47	97.9	15	2 AAW38924	Aaw38924 Peptide r
27	47	97.9	15	7 ADB49285	Adb49285 Novel WW
28	47	97.9	17	3 AAB17217	Aab17217 Masc cell
29	47	97.9	17	5 ABB73208	Abb73208 Mast cell
30	47	97.9	17	7 ADJ73362	Adj73362 Mast cell
31	47	97.9	17	8 ADJ52996	Adj52996 CH1 delet
32	47	97.9	17	8 ADJ51957	Adj51957 CH1 delet
33	47	97.9	22	2 AAY41645	Aay41645 Human pep
34	47	97.9	25	2 AAY41612	Aay41612 Mammalian
35	47	97.9	613	6 ABR42419	AbR42419 Human pot
36	46	95.8	12	2 AAR93362	Aar93362 LYN/P13K
37	46	95.8	12	2 AAR93377	Aar93377 Grb-2 pro
38	46	95.8	12	2 AAR93345	Aar93345 SRC prote
39	46	95.8	12	3 AAB17255	Aab17255 SH3 antag
40	46	95.8	12	5 ABB73248	Abb73248 Src homol
41	46	95.8	12	7 ADJ73402	Adj73402 SH3 antag
42	46	95.8	12	8 ADJ53036	Adj53036 CH1 delet
43	46	95.8	12	8 ADJ51997	Adj51997 CH1 delet
44	46	95.8	14	2 AAR93386	Aar93386 GST-SRC S
45	46	95.8	14	2 AAR93404	Aar93404 GST-FYN S
46	46	95.8	14	2 AAR93450	Aar93450 GST-P13K
47	46	95.8	14	2 AAR93388	Aar93388 GST-SRC S
48	46	95.8	14	2 AAR93392	Aar93392 GST-FYN S
49	46	95.8	14	2 AAR93400	Aar93400 GST-FYN S
50	46	95.8	14	2 AAR93428	Aar93428 GST-YES S
51	46	95.8	14	2 AAR93434	Aar93434 GST-YES S
52	46	95.8	14	2 AAR93439	Aar93439 GST-YES S
53	46	95.8	14	2 AAR93381	Aar93381 GST-SRC S
54	46	95.8	14	2 AAR93383	Aar93383 GST-SRC S
55	46	95.8	14	2 AAR93410	Aar93410 GST-LYN S
56	46	95.8	14	2 AAR93452	Aar93452 GST-P13K
57	46	95.8	14	2 AAR93396	Aar93396 GST-FYN S
58	46	95.8	14	2 AAR93402	Aar93402 GST-FYN S
59	46	95.8	14	2 AAR93427	Aar93427 GST-YES S
60	46	95.8	14	2 AAR93432	Aar93432 GST-YES S
61	46	95.8	14	2 AAR93385	Aar93385 GST-SRC S
62	46	95.8	14	2 AAR93419	Aar93419 GST-LYN S
63	46	95.8	14	2 AAR93420	Aar93420 GST-LYN S
64	46	95.8	14	2 AAR93431	Aar93431 GST-YES S
65	46	95.8	14	2 AAR93446	Aar93446 GST-P13K
66	46	95.8	14	2 AAR93390	Aar93390 GST-SRC S
67	46	95.8	14	2 AAR93399	Aar93399 GST-FYN S
68	46	95.8	14	2 AAR93406	Aar93406 GST-FYN S
69	46	95.8	14	2 AAR93411	Aar93411 GST-LYN S
70	46	95.8	14	2 AAR93451	Aar93451 GST-P13K
71	46	95.8	14	2 AAR93455	Aar93455 GST-P13K
72	46	95.8	14	2 AAR93401	Aar93401 GST-FYN S
73	46	95.8	14	2 AAR93414	Aar93414 GST-LYN S
74	46	95.8	14	2 AAR93418	Aar93418 GST-LYN S
75	46	95.8	14	2 AAR93389	Aar93389 GST-SRC S
76	46	95.8	14	2 AAR93405	Aar93405 GST-FYN S
77	46	95.8	14	2 AAR93433	Aar93433 GST-YES S
78	46	95.8	14	2 AAR93444	Aar93444 GST-P13K
79	46	95.8	14	2 AAR93403	Aar93403 GST-FYN S
80	46	95.8	14	2 AAR93424	Aar93424 GST-LYN S
81	46	95.8	14	2 AAR93387	Aar93387 GST-SRC S
82	46	95.8	14	2 AAR93412	Aar93412 GST-LYN S
83	46	95.8	14	2 AAR93425	Aar93425 GST-YES S
84	46	95.8	14	2 AAR93441	Aar93441 GST-YES S
85	46	95.8	14	2 AAR93454	Aar93454 GST-P13K
86	46	95.8	14	2 AAR93393	Aar93393 GST-FYN S
87	46	95.8	14	2 AAR93398	Aar93398 GST-FYN S
88	46	95.8	14	2 AAR93413	Aar93413 GST-LYN S
89	46	95.8	14	2 AAR93416	Aar93416 GST-LYN S
90	46	95.8	14	2 AAR93422	Aar93422 GST-LYN S
91	46	95.8	14	2 AAR93436	Aar93436 GST-YES S
92	46	95.8	14	2 AAR93442	Aar93442 GST-YES S
93	46	95.8	14	2 AAR93445	Aar93445 GST-P13K
94	46	95.8	14	2 AAR93391	Aar93391 GST-FYN S
95	46	95.8	14	2 AAR93394	Aar93394 GST-FYN S
96	46	95.8	14	2 AAR93417	Aar93417 GST-LYN S
97	46	95.8	14	2 AAR93429	Aar93429 GST-YES S

98	46	95.8	14	2	AAR93430	Aar93430	GST-YES S	171	45	93.8	85	6	ABO02918	Human sec
99	46	95.8	14	2	AAR93435	Aar93435	GST-YES S	172	45	93.8	85	6	ABR75072	Human sec
100	46	95.8	14	2	AAR93437	Aar93437	GST-YES S	173	45	93.8	85	6	ABR94834	Human sec
101	46	95.8	14	2	AAR93438	Aar93438	GST-YES S	174	45	93.8	85	6	ABO25164	Novel hum
102	46	95.8	14	2	AAR93440	Aar93440	GST-YES S	175	45	93.8	85	6	ABO33760	Novel hum
103	46	95.8	14	2	AAR93397	Aar93397	GST-FYN S	176	45	93.8	85	6	ABU85807	Human PRO
104	46	95.8	14	2	AAR93407	Aar93407	GST-FYN S	177	45	93.8	85	6	ABU98967	Novel hum
105	46	95.8	14	2	AAR93447	Aar93447	GST-PI3K	178	45	93.8	85	6	ABU98182	Novel hum
106	46	95.8	14	2	AAR93453	Aar93453	GST-PI3K	179	45	93.8	85	6	ABU91888	Novel hum
107	46	95.8	14	2	AAR93456	Aar93456	GST-PI3K	180	45	93.8	85	6	ABU89581	Human PRO
108	46	95.8	14	2	AAR93426	Aar93426	GST-YES S	181	45	93.8	85	6	ABU86422	Human sec
109	46	95.8	14	2	AAR93443	Aar93443	GST-YES S	182	45	93.8	85	6	ABU67635	Human sec
110	46	95.8	14	2	AAR93448	Aar93448	GST-PI3K	183	45	93.8	85	6	ABU80663	Human PRO
111	46	95.8	14	2	AAR93382	Aar93382	GST-SRC S	184	45	93.8	85	6	ABR99581	Human sec
112	46	95.8	14	2	AAR93384	Aar93384	GST-SRC S	185	45	93.8	85	6	ABR98971	Human sec
113	46	95.8	14	2	AAR93408	Aar93408	GST-LYN S	186	45	93.8	85	6	ABO16494	Human sec
114	46	95.8	14	2	AAR93409	Aar93409	GST-LYN S	187	45	93.8	85	6	ABR92394	Human sec
115	46	95.8	14	2	AAR93421	Aar93421	GST-LYN S	188	45	93.8	85	6	ABO19035	Human sec
116	46	95.8	14	2	AAR93449	Aar93449	GST-PI3K	189	45	93.8	85	6	ABR78456	Human sec
117	46	95.8	14	2	AAR93395	Aar93395	GST-FYN S	190	45	93.8	85	6	ABU85192	Novel hum
118	46	95.8	14	2	AAR93415	Aar93415	GST-LYN S	191	45	93.8	85	6	ABO00331	Novel hum
119	46	95.8	14	2	AAR93423	Aar93423	GST-LYN S	192	45	93.8	85	6	ABO11663	Human sec
120	46	95.8	202	4	ABG52649	ABG52649	Porcine F	193	45	93.8	85	6	ABO02308	Human sec
121	45	93.8	45	2	AAI36531	AAI36531	Fragment	194	45	93.8	85	6	ADA40702	Human sec
122	45	93.8	45	6	ADA12079	ADA12079	Human nov	195	45	93.8	85	6	ABU88882	Novel hum
123	45	93.8	68	4	ABBS1247	ABBS1247	Human ami	196	45	93.8	85	6	ABU67282	Novel hum
124	45	93.8	78	2	AAI12633	AAI12633	Human 5'	197	45	93.8	85	6	ABU83577	Human sec
125	45	93.8	83	2	AAI12632	AAI12632	Human 5'	198	45	93.8	85	6	ABO06378	Novel hum
126	45	93.8	83	4	ABBS0241	ABBS0241	Human bre	199	45	93.8	85	6	ABR59414	Human sec
127	45	93.8	83	4	AAU33342	AAU33342	Human bre	200	45	93.8	85	6	ABO09476	Human sec
128	45	93.8	83	5	ABGG6759	ABGG6759	Human pro	201	45	93.8	85	6	ABO19340	Novel hum
129	45	93.8	83	5	AAO15776	AAO15776	Human bre	202	45	93.8	85	6	ABO11358	Human sec
130	45	93.8	83	5	ABG78909	ABG78909	Human bre	203	45	93.8	85	6	ABR66976	Human sec
131	45	93.8	83	6	ABJ37732	ABJ37732	Human tum	204	45	93.8	85	6	ABO16189	Human sec
132	45	93.8	83	7	ADL52724	ADL52724	Human bre	205	45	93.8	85	6	ABO13895	Human sec
133	45	93.8	83	8	ADL44014	ADL44014	Human bre	206	45	93.8	85	6	ABU65798	Human sec
134	45	93.8	85	2	AAWS9652	AAWS9652	Amino aci	207	45	93.8	85	6	ABO07646	Human PRO
135	45	93.8	85	2	AAW04146	AAW04146	Human 5'	208	45	93.8	85	6	ABO03833	Human sec
136	45	93.8	85	2	AAW93615	AAW93615	Human 5'	209	45	93.8	85	6	ABR67281	Human sec
137	45	93.8	85	2	AAW04164	AAW04164	Human 5'	210	45	93.8	85	6	ABR57281	Human sec
138	45	93.8	85	2	AAI36038	AAI36038	Extended	211	45	93.8	85	6	ABO15884	Human sec
139	45	93.8	85	2	AAI36037	AAI36037	Extended	212	45	93.8	85	6	ABU56165	Human sec
140	45	93.8	85	2	AAI35880	AAI35880	Extended	213	45	93.8	85	6	ABU65493	Human PRO
141	45	93.8	85	2	AAI59692	AAI59692	Secreted	214	45	93.8	85	6	ABU95438	Novel hum
142	45	93.8	85	2	AAI59643	AAI59643	Secreted	215	45	93.8	85	6	ABU71341	Human PRO
143	45	93.8	85	2	AAI01589	AAI01589	Secreted	216	45	93.8	85	6	ABO07951	Human PRO
144	45	93.8	85	2	AAI12981	AAI12981	Human sec	217	45	93.8	85	6	ABR70192	Human sec
145	45	93.8	85	2	AAI25449	AAI25449	Human sec	218	45	93.8	85	6	ABR69525	Human sec
146	45	93.8	85	2	AAI12670	AAI12670	Human 5'	219	45	93.8	85	6	ABO01666	Human PRO
147	45	93.8	85	3	AAO00012	AAO00012	Human sec	220	45	93.8	85	6	ABR60265	Human sec
148	45	93.8	85	4	AAI31193	AAI31193	Amino aci	221	45	93.8	85	6	ABR68000	Human sec
149	45	93.8	85	4	AAU29253	AAU29253	Human PRO	222	45	93.8	85	6	ABR65388	Human sec
150	45	93.8	85	4	AAI31245	AAI31245	Human PRO	223	45	93.8	85	6	ABR68610	Human sec
151	45	93.8	85	4	AAI30699	AAI30699	Human CH3	224	45	93.8	85	6	ABR72022	Human sec
152	45	93.8	85	5	AAU83647	AAU83647	Human PRO	225	45	93.8	85	6	ABU85502	Human PRO
153	45	93.8	85	5	ABBS4934	ABBS4934	Human PRO	226	45	93.8	85	6	ABU89192	Human sec
154	45	93.8	85	5	AAE20461	AAE20461	Human tum	227	45	93.8	85	6	ABU83272	Human sec
155	45	93.8	85	5	ABR95540	ABR95540	Human ang	228	45	93.8	85	6	ABU95128	Novel hum
156	45	93.8	85	5	ADY31852	ADY31852	Novel hum	229	45	93.8	85	6	ABU90676	Novel hum
157	45	93.8	85	6	ABU58629	ABU58629	Human PRO	230	45	93.8	85	6	ABU84187	Human sec
158	45	93.8	85	6	ABU88177	ABU88177	Novel hum	231	45	93.8	85	6	ABU93838	Novel hum
159	45	93.8	85	6	ABU84492	ABU84492	Human sec	232	45	93.8	85	6	ABR65083	Human sec
160	45	93.8	85	6	ABR66366	ABR66366	Human sec	233	45	93.8	85	6	ABR68915	Human sec
161	45	93.8	85	6	ABR5756	ABR5756	Human sec	234	45	93.8	85	6	ABO06731	Human sec
162	45	93.8	85	6	ABU99696	ABU99696	Human sec	235	45	93.8	85	6	ABR99276	Human sec
163	45	93.8	85	6	ABU82935	ABU82935	Human PRO	236	45	93.8	85	6	ABU57160	Human PRO
164	45	93.8	85	6	ABU90056	ABU90056	Novel hum	237	45	93.8	85	6	ABU72050	Novel hum
165	45	93.8	85	6	ABR68305	ABR68305	Human sec	238	45	93.8	85	6	ABU86112	Novel hum
166	45	93.8	85	6	ADA56852	ADA56852	Human sec	239	45	93.8	85	6	ABU82399	Novel hum
167	45	93.8	85	6	ABU96388	ABU96388	Novel hum	240	45	93.8	85	6	ABU87410	Human PRO
168	45	93.8	85	6	ABU92789	ABU92789	Human sec	241	45	93.8	85	6	ABU67151	Novel hum
169	45	93.8	85	6	ABU80794	ABU80794	Human PRO	242	45	93.8	85	6	ABU83882	Human sec
170	45	93.8	85	6	ABO08866	ABO08866	Human sec	243	45	93.8	85	6	ABO08256	Human PRO

244	45	93.8	85	6	ABU81967	Novel hum	317	45	93.8	85	6	ABO48262	Human sec
245	45	93.8	85	6	ABU66131	Novel hum	318	45	93.8	85	6	ABR93004	Human sec
246	45	93.8	85	6	ABR5960	Human sec	319	45	93.8	85	6	ABO24765	Human sec
247	45	93.8	85	6	ABU94148	Novel hum	320	45	93.8	85	6	ABM11776	Human sec
248	45	93.8	85	6	ABO00021	Novel hum	321	45	93.8	85	6	ABM02877	Human sec
249	45	93.8	85	6	ABU82103	Novel hum	322	45	93.8	85	6	ABM16173	Human sec
250	45	93.8	85	6	ABR66671	Human sec	323	45	93.8	85	6	ABO27734	Human sec
251	45	93.8	85	6	ABR91089	Human sec	324	45	93.8	85	6	ABM29225	Human sec
252	45	93.8	85	6	ABG75947	Human ant	325	45	93.8	85	6	ABM07201	Human sec
253	45	93.8	85	6	ABU94516	Human PRO	326	45	93.8	85	6	ABM21295	Human sec
254	45	93.8	85	6	ABU79398	Human PRO	327	45	93.8	85	6	ABM09641	Human sec
255	45	93.8	85	6	ABU86727	Human sec	328	45	93.8	85	6	ABO41511	Human sec
256	45	93.8	85	6	ABU87032	Novel hum	329	45	93.8	85	6	ABO36326	Human PRO
257	45	93.8	85	6	ABU94821	Human PRO	330	45	93.8	85	6	ABO43855	Human PRO
258	45	93.8	85	6	ABO04748	Human PRO	331	45	93.8	85	6	ABM76555	Human sec
259	45	93.8	85	6	ABR70497	Human sec	332	45	93.8	85	6	ABM76251	Human sec
260	45	93.8	85	6	ABU98662	Human PRO	333	45	93.8	85	6	ABM25870	Human sec
261	45	93.8	85	6	ABR66061	Human sec	334	45	93.8	85	6	ABM26175	Human sec
262	45	93.8	85	6	ABR64778	Human sec	335	45	93.8	85	6	ABO03528	Human sec
263	45	93.8	85	6	ABU79703	Human PRO	336	45	93.8	85	6	ABO02613	Human sec
264	45	93.8	85	6	ABU93094	Human sec	337	45	93.8	85	6	ABR90784	Human sec
265	45	93.8	85	6	ABU96053	Human PRO	338	45	93.8	85	6	ABR73852	Human sec
266	45	93.8	85	6	ABU91273	Novel hum	339	45	93.8	85	6	ABO17104	Human sec
267	45	93.8	85	6	ABU90366	Novel hum	340	45	93.8	85	6	ABR94529	Human sec
268	45	93.8	85	6	ABO09781	Human sec	341	45	93.8	85	6	ABR76036	Human sec
269	45	93.8	85	6	ABO11053	Human sec	342	45	93.8	85	6	ABR71412	Human sec
270	45	93.8	85	6	ABR71107	Human sec	343	45	93.8	85	6	ABR93309	Human sec
271	45	93.8	85	6	ABU87715	Human PRO	344	45	93.8	85	6	ABR93614	Human sec
272	45	93.8	85	6	ABU91583	Human PRO	345	45	93.8	85	6	ABR88039	Human sec
273	45	93.8	85	6	ABU84797	Human sec	346	45	93.8	85	6	ABO28039	Human sec
274	45	93.8	85	6	ABR69887	Human sec	347	45	93.8	85	6	ABO30174	Human sec
275	45	93.8	85	6	ABU80264	Human PRO	348	45	93.8	85	6	ABO33383	Human PRO
276	45	93.8	85	6	ABU93533	Human PRO	349	45	93.8	85	6	ABM05071	Human sec
277	45	93.8	85	6	ABO10086	Human sec	350	45	93.8	85	6	ABM09031	Human sec
278	45	93.8	85	6	ABO09171	Human sec	351	45	93.8	85	6	ABO36631	Human sec
279	45	93.8	85	6	ABU10739	Human sec	352	45	93.8	85	6	ABO35716	Human PRO
280	45	93.8	85	6	ABU95748	Human PRO	353	45	93.8	85	6	ABO39681	Human sec
281	45	93.8	85	6	ABU79793	Human sec	354	45	93.8	85	6	ABM10556	Human sec
282	45	93.8	85	6	ABU96957	Novel hum	355	45	93.8	85	6	ABM12081	Human sec
283	45	93.8	85	6	ABR70802	Human sec	356	45	93.8	85	6	ABO52227	Human PRO
284	45	93.8	85	6	ABO05153	Novel hum	357	45	93.8	85	6	ABO52532	Human PRO
285	45	93.8	85	6	ABO08561	Human sec	358	45	93.8	85	6	ABO23850	Human sec
286	45	93.8	85	6	ABO05768	Human sec	359	45	93.8	85	6	ABR97336	Human sec
287	45	93.8	85	6	ABR74157	Human sec	360	45	93.8	85	6	ABR87124	Human sec
288	45	93.8	85	6	ABR95749	Human sec	361	45	93.8	85	6	ABM11166	Human sec
289	45	93.8	85	6	ABR81046	Human sec	362	45	93.8	85	6	ABM28310	Human sec
290	45	93.8	85	6	ABR81351	Human sec	363	45	93.8	85	6	ABO32309	Human sec
291	45	93.8	85	6	ABM01047	Human sec	364	45	93.8	85	6	ABM15436	Human sec
292	45	93.8	85	6	ABR88649	Human sec	365	45	93.8	85	6	ABM06591	Human sec
293	45	93.8	85	6	ABM77470	Human sec	366	45	93.8	85	6	ABM04402	Human sec
294	45	93.8	85	6	ABO28954	Human sec	367	45	93.8	85	6	ABM22515	Human sec
295	45	93.8	85	6	ABO31699	Human sec	368	45	93.8	85	6	ABM07811	Human sec
296	45	93.8	85	6	ABM08116	Human sec	369	45	93.8	85	6	ABO40901	Human sec
297	45	93.8	85	6	ABO40596	Human sec	370	45	93.8	85	6	ABM35548	Human sec
298	45	93.8	85	6	ABO36021	Human PRO	371	45	93.8	85	6	ABM33311	Human sec
299	45	93.8	85	6	ABO44160	Human PRO	372	45	93.8	85	6	ABO52837	Human PRO
300	45	93.8	85	6	ADA78212	Human sec	373	45	93.8	85	6	ABO50397	Human sec
301	45	93.8	85	6	ABM24955	Human sec	374	45	93.8	85	6	ABU99391	Human sec
302	45	93.8	85	6	ABO03223	Human sec	375	45	93.8	85	6	ABO04443	Human sec
303	45	93.8	85	6	ABR90479	Human sec	376	45	93.8	85	6	ABO06073	Human sec
304	45	93.8	85	6	ABM17393	Human sec	377	45	93.8	85	6	ABM18613	Human sec
305	45	93.8	85	6	ABR95139	Human sec	378	45	93.8	85	6	ABR97641	Human sec
306	45	93.8	85	6	ABR95444	Human sec	379	45	93.8	85	6	ABR80741	Human sec
307	45	93.8	85	6	ABJ72283	Human PRO	380	45	93.8	85	6	ABM01352	Human sec
308	45	93.8	85	6	ABO21682	Human sec	381	45	93.8	85	6	ABR88954	Human sec
309	45	93.8	85	6	ABR97946	Human sec	382	45	93.8	85	6	ABM13606	Human sec
310	45	93.8	85	6	ABR87734	Human sec	383	45	93.8	85	6	ABM20990	Human sec
311	45	93.8	85	6	ABM77775	Human sec	384	45	93.8	85	6	ABO42121	Human sec
312	45	93.8	85	6	ABM28005	Human sec	385	45	93.8	85	6	ABO42731	Human sec
313	45	93.8	85	6	ABM06286	Human sec	386	45	93.8	85	6	ABM10251	Human sec
314	45	93.8	85	6	ABM03792	Human sec	387	45	93.8	85	6	ABO38766	Human sec
315	45	93.8	85	6	ABM35243	Human sec	388	45	93.8	85	6	ABM33006	Human sec
316	45	93.8	85	6	ABM26480	Human sec	389	45	93.8	85	6	ABM22820	Human sec

390	45	93.8	85	6	ABM75031	Human sec	463	45	93.8	85	6	ABM25565	Human sec
391	45	93.8	85	6	ADA80004	Human sec	464	45	93.8	85	6	ABM20075	Human sec
392	45	93.8	85	6	ABR96421	Human sec	465	45	93.8	85	6	ABO46981	Human PRO
393	45	93.8	85	6	ABM02572	Human sec	466	45	93.8	85	6	ABO47286	Human PRO
394	45	93.8	85	6	ABR86514	Human sec	467	45	93.8	85	6	ADA83529	Human sec
395	45	93.8	85	6	ABR86819	Human sec	468	45	93.8	85	6	ABR71717	Human sec
396	45	93.8	85	6	ABM16783	Human sec	469	45	93.8	85	6	ABR72327	Human sec
397	45	93.8	85	6	ABM29835	Human sec	470	45	93.8	85	6	ABR98666	Human sec
398	45	93.8	85	6	ABO29259	Human sec	471	45	93.8	85	6	ABO07036	Human sec
399	45	93.8	85	6	ABM24040	Human sec	472	45	93.8	85	6	ABR84989	Human sec
400	45	93.8	85	6	ABM23430	Human sec	473	45	93.8	85	6	ABR73547	Human sec
401	45	93.8	85	6	ADA47228	Human sec	474	45	93.8	85	6	ABR76641	Human sec
402	45	93.8	85	6	ABM22210	Human sec	475	45	93.8	85	6	ABR73242	Human sec
403	45	93.8	85	6	ABO37851	Human sec	476	45	93.8	85	6	ABJ72411	Human PRO
404	45	93.8	85	6	ABM28615	Human sec	477	45	93.8	85	6	ABM18308	Human sec
405	45	93.8	85	6	ABM28920	Human sec	478	45	93.8	85	6	ABO20767	Human sec
406	45	93.8	85	6	ABM65664	Human sec	479	45	93.8	85	6	ABO25510	Human PRO
407	45	93.8	85	6	ABM75946	Human sec	480	45	93.8	85	6	ABO25815	Human PRO
408	45	93.8	85	6	ABM34226	Human sec	481	45	93.8	85	6	ABR94224	Human sec
409	45	93.8	85	6	ABM34531	Human sec	482	45	93.8	85	6	ABR80131	Human sec
410	45	93.8	85	6	ABO20462	Human sec	483	45	93.8	85	6	ABM11471	Human sec
411	45	93.8	85	6	ABO21377	Human sec	484	45	93.8	85	6	ABO33078	Human PRO
412	45	93.8	85	6	ABO22292	Human sec	485	45	93.8	85	6	ABO30784	Human sec
413	45	93.8	85	6	ABR96726	Human sec	486	45	93.8	85	6	ABO31089	Human sec
414	45	93.8	85	6	ABR85904	Human sec	487	45	93.8	85	6	ABM27395	Human sec
415	45	93.8	85	6	ABR99886	Human sec	488	45	93.8	85	6	ABM30140	Human sec
416	45	93.8	85	6	ABM00742	Human sec	489	45	93.8	85	6	ABM05676	Human sec
417	45	93.8	85	6	ABM00437	Human sec	490	45	93.8	85	6	ABM15741	Human sec
418	45	93.8	85	6	ABO29869	Human sec	491	45	93.8	85	6	ABM08726	Human sec
419	45	93.8	85	6	ABM23735	Human sec	492	45	93.8	85	6	ABO42426	Human sec
420	45	93.8	85	6	ABM29530	Human sec	493	45	93.8	85	6	ABO38156	Human sec
421	45	93.8	85	6	ABO38461	Human sec	494	45	93.8	85	6	ABO46066	Human PRO
422	45	93.8	85	6	ABO45761	Human PRO	495	45	93.8	85	6	ABM66869	Human sec
423	45	93.8	85	6	ABM20685	Human sec	496	45	93.8	85	6	ABR20572	Human sec
424	45	93.8	85	6	ADA81731	Human sec	497	45	93.8	85	6	ABM19770	Human sec
425	45	93.8	85	6	ABO16799	Human sec	498	45	93.8	85	6	ABO49482	Human sec
426	45	93.8	85	6	ABO18425	Human sec	499	45	93.8	85	6	ABO49787	Human sec
427	45	93.8	85	6	ABO22852	Human PRO	500	45	93.8	85	6	ADA78824	Human sec
428	45	93.8	85	6	ABO23157	Human PRO	501	45	93.8	85	6	ABO34306	Human sec
429	45	93.8	85	6	ABR92699	Human sec	502	45	93.8	85	6	ABR88344	Human sec
430	45	93.8	85	6	ABR81656	Human sec	503	45	93.8	85	6	ABM27090	Human sec
431	45	93.8	85	6	ABM78080	Human sec	504	45	93.8	85	6	ABM03487	Human sec
432	45	93.8	85	6	ABR89869	Human sec	505	45	93.8	85	6	ABO39986	Human sec
433	45	93.8	85	6	ABM26785	Human sec	506	45	93.8	85	7	ABO50092	Human sec
434	45	93.8	85	6	ABM13911	Human sec	507	45	93.8	85	7	ABO51007	Human sec
435	45	93.8	85	6	ABO28649	Human sec	508	45	93.8	85	7	ABO05463	Human sec
436	45	93.8	85	6	ABO30479	Human sec	509	45	93.8	85	7	ABR74767	Human sec
437	45	93.8	85	6	ABM07506	Human sec	510	45	93.8	85	7	ABR77246	Human sec
438	45	93.8	85	6	ABM04097	Human sec	511	45	93.8	85	7	ABM18003	Human sec
439	45	93.8	85	6	ABO37241	Human sec	512	45	93.8	85	7	ABR96054	Human sec
440	45	93.8	85	6	ABO41816	Human sec	513	45	93.8	85	7	ABO21987	Human sec
441	45	93.8	85	6	ABO35411	Human PRO	514	45	93.8	85	7	ABO20157	Human sec
442	45	93.8	85	6	ABM25260	Human sec	515	45	93.8	85	7	ABO24460	Human sec
443	45	93.8	85	6	ABO47652	Human sec	516	45	93.8	85	7	ABR86209	Human sec
444	45	93.8	85	6	ABO47957	Human sec	517	45	93.8	85	7	ABM10861	Human sec
445	45	93.8	85	6	ABO48567	Human sec	518	45	93.8	85	7	ABM76860	Human sec
446	45	93.8	85	6	ABO51617	Human PRO	519	45	93.8	85	7	ABR99564	Human sec
447	45	93.8	85	6	ABO51922	Human PRO	520	45	93.8	85	7	ABM12691	Human sec
448	45	93.8	85	6	ABO50702	Human sec	521	45	93.8	85	7	ABM05981	Human sec
449	45	93.8	85	6	ABR79826	Human sec	522	45	93.8	85	7	ABO35106	Human PRO
450	45	93.8	85	6	ABM17088	Human sec	523	45	93.8	85	7	ABM03182	Human sec
451	45	93.8	85	6	ABO18120	Human sec	524	45	93.8	85	7	ABM19160	Human sec
452	45	93.8	85	6	ABO21072	Human sec	525	45	93.8	85	7	ABM19465	Human sec
453	45	93.8	85	6	ABR97031	Human sec	526	45	93.8	85	7	ABO46676	Human PRO
454	45	93.8	85	6	ABM12386	Human sec	527	45	93.8	85	7	ABO49177	Human sec
455	45	93.8	85	6	ABM16478	Human sec	528	45	93.8	85	7	ABR69220	Human sec
456	45	93.8	85	6	ABM24345	Human sec	529	45	93.8	85	7	ABR89259	Human sec
457	45	93.8	85	6	ABM14826	Human sec	530	45	93.8	85	7	ABR72632	Human sec
458	45	93.8	85	6	ABM04707	Human sec	531	45	93.8	85	7	ABR74462	Human sec
459	45	93.8	85	6	ABM06896	Human sec	532	45	93.8	85	7	ABO18730	Human sec
460	45	93.8	85	6	ABM09336	Human sec	533	45	93.8	85	7	ABR80436	Human sec
461	45	93.8	85	6	ABO39376	Human sec	534	45	93.8	85	7	ABM01657	Human sec
462	45	93.8	85	6	ABM75641	Human sec	535	45	93.8	85	7	ABM02267	Human sec

536 45 93.8 85 7 ABR87429 Human sec
537 45 93.8 85 7 ABM12996 Human sec
538 45 93.8 85 7 ABM30750 Human sec
539 45 93.8 85 7 ABM24650 Human sec
540 45 93.8 85 7 ABO29564 Human sec
541 45 93.8 85 7 ABO311394 Human sec
542 45 93.8 85 7 ABM14521 Human sec
543 45 93.8 85 7 ABM09946 Human sec
544 45 93.8 85 7 ABO39071 Human sec
545 45 93.8 85 7 ABM34836 Human sec
546 45 93.8 85 7 ABO51312 Human sec
547 45 93.8 85 7 ABO04138 Human sec
548 45 93.8 85 7 ABO10608 Human PRO
549 45 93.8 85 7 ABR77851 Human sec
550 45 93.8 85 7 ABR79061 Human sec
551 45 93.8 85 7 ABO24155 Human sec
552 45 93.8 85 7 ABR93919 Human sec
553 45 93.8 85 7 ABM01962 Human sec
554 45 93.8 85 7 ABM78385 Human sec
555 45 93.8 85 7 ABR90174 Human sec
556 45 93.8 85 7 ABM27700 Human sec
557 45 93.8 85 7 ABM13301 Human sec
558 45 93.8 85 7 ABO32004 Human sec
559 45 93.8 85 7 ABM14216 Human sec
560 45 93.8 85 7 ABM08421 Human sec
561 45 93.8 85 7 ABO40291 Human sec
562 45 93.8 85 7 ABO47426 Human sec
563 45 93.8 85 7 ABM33921 Human sec
564 45 93.8 85 7 ABM20380 Human sec
565 45 93.8 85 7 ABO48872 Human sec
566 45 93.8 85 7 ABR72937 Human sec
567 45 93.8 85 7 ABO15579 Human sec
568 45 93.8 85 7 ABR85294 Human sec
569 45 93.8 85 7 ABO15274 Human sec
570 45 93.8 85 7 ABO17409 Human sec
571 45 93.8 85 7 ABO19851 Human sec
572 45 93.8 85 7 ABM17698 Human sec
573 45 93.8 85 7 ABJ72113 Human mem
574 45 93.8 85 7 ABR85599 Human sec
575 45 93.8 85 7 ABM77165 Human sec
576 45 93.8 85 7 ABO28344 Human sec
577 45 93.8 85 7 ABM23125 Human sec
578 45 93.8 85 7 ABM30445 Human sec
579 45 93.8 85 7 ABM21905 Human sec
580 45 93.8 85 7 ABM21600 Human sec
581 45 93.8 85 7 ABM15131 Human sec
582 45 93.8 85 7 ABO41206 Human sec
583 45 93.8 85 7 ABO36936 Human sec
584 45 93.8 85 7 ABO37546 Human sec
585 45 93.8 85 7 ABM75336 Human sec
586 45 93.8 85 7 ABM33616 Human sec
587 45 93.8 85 7 ABO46371 Human PRO
588 45 93.8 85 7 ADA82895 Human sec
589 45 93.8 85 7 ADB83602 Novel hum
590 45 93.8 85 7 ADB80708 Novel hum
591 45 93.8 85 7 ADB73249 Novel hum
592 45 93.8 85 7 ABM31970 Human sec
593 45 93.8 85 7 ADB78331 Novel hum
594 45 93.8 85 7 ABM31360 Human sec
595 45 93.8 85 7 ADB84979 Human PRO
596 45 93.8 85 7 ADB78085 Human sec
597 45 93.8 85 7 ADB86203 Human sec
598 45 93.8 85 7 ABM32275 Human sec
599 45 93.8 85 7 ABM32580 Human PRO
600 45 93.8 85 7 ADB87151 Human PRO
601 45 93.8 85 7 ADB84733 Human PRO
602 45 93.8 85 7 ABM31665 Human sec
603 45 93.8 85 7 ABM31055 Human sec
604 45 93.8 85 7 ADB83848 Novel hum
605 45 93.8 85 7 ADB73003 Novel hum
606 45 93.8 85 7 ADC36841 Human PRO
607 45 93.8 85 7 ADC21831 Human PRO
608 45 93.8 85 7 ADC49862 Novel hum

609 45 93.8 85 7 ADC49061 Novel hum
610 45 93.8 85 7 ADC49578 Novel hum
611 45 93.8 85 7 ADC47439 Novel hum
612 45 93.8 85 7 ADC47184 Novel hum
613 45 93.8 85 7 ADC78059 Novel hum
614 45 93.8 85 7 ADD06294 Novel hum
615 45 93.8 85 7 ADD05933 Human sec
616 45 93.8 85 7 ADD10525 Human sec
617 45 93.8 85 7 ADD77813 Novel hum
618 45 93.8 85 7 ADD11485 Human sec
619 45 93.8 85 7 ADD50776 Novel hum
620 45 93.8 85 7 ADD51022 Novel hum
621 45 93.8 85 7 ADD37278 Human sec
622 45 93.8 85 7 ADD50503 Human PRO
623 45 93.8 85 7 ADD50257 Human PRO
624 45 93.8 85 7 ADD51268 Novel hum
625 45 93.8 85 7 ADG02928 Novel hum
626 45 93.8 85 7 ADG01635 Novel hum
627 45 93.8 85 7 ADF95810 Novel hum
628 45 93.8 85 7 ADG12625 Novel hum
629 45 93.8 85 7 ADH09285 Human PRO
630 45 93.8 85 7 ADG63737 Human sec
631 45 93.8 85 7 ADL33064 Novel hum
632 45 93.8 85 7 ADM30600 Novel hum
633 45 93.8 85 7 ADM77842 Human PRO
634 45 93.8 85 7 ADC48815 Novel hum
635 45 93.8 85 8 ADE20986 Novel hum
636 45 93.8 85 8 ADE05830 Human PRO
637 45 93.8 85 8 ADD75059 Human PRO
638 45 93.8 85 8 ADD75805 Novel hum
639 45 93.8 85 8 ADD85037 Novel hum
640 45 93.8 85 8 ADD86863 Novel hum
641 45 93.8 85 8 ADE20740 Novel hum
642 45 93.8 85 8 ADE39037 Novel hum
643 45 93.8 85 8 ADE05584 Human PRO
644 45 93.8 85 8 ADD73569 Human PRO
645 45 93.8 85 8 ADD78409 Novel hum
646 45 93.8 85 8 ADE41486 Human sec
647 45 93.8 85 8 ADE74597 Human sec
648 45 93.8 85 8 ADE21232 Novel hum
649 45 93.8 85 8 ADE77347 Novel hum
650 45 93.8 85 8 ADE20494 Novel hum
651 45 93.8 85 8 ADD75559 Human PRO
652 45 93.8 85 8 ADD74075 Human PRO
653 45 93.8 85 8 ADD74321 Human PRO
654 45 93.8 85 8 ADD76051 Novel hum
655 45 93.8 85 8 ADD85543 Novel hum
656 45 93.8 85 8 ADE05092 Human PRO
657 45 93.8 85 8 ADD75305 Human PRO
658 45 93.8 85 8 ADD76849 Novel hum
659 45 93.8 85 8 ADD86617 Novel hum
660 45 93.8 85 8 ADE41151 Human sec
661 45 93.8 85 8 ADD78085 Novel hum
662 45 93.8 85 8 ADE75209 Human sec
663 45 93.8 85 8 ADD77593 Novel hum
664 45 93.8 85 8 ADD77839 Novel hum
665 45 93.8 85 8 ADD85297 Novel hum
666 45 93.8 85 8 ADD73829 Human PRO
667 45 93.8 85 8 ADD74567 Human PRO
668 45 93.8 85 8 ADD77095 Novel hum
669 45 93.8 85 8 ADD85789 Novel hum
670 45 93.8 85 8 ADE05338 Human PRO
671 45 93.8 85 8 ADD74813 Human PRO
672 45 93.8 85 8 AG05625 Novel hum
673 45 93.8 85 8 AG27179 Human PRO
674 45 93.8 85 8 ADF96422 Novel hum
675 45 93.8 85 8 ADG11242 Novel hum
676 45 93.8 85 8 AG04693 Novel hum
677 45 93.8 85 8 ADG12021 Novel hum
678 45 93.8 85 8 ADG00853 Novel hum
679 45 93.8 85 8 ADF94578 Novel hum
680 45 93.8 85 8 ADG06674 Human PRO
681 45 93.8 85 8 ADG83109 Human PRO

682	45	93.8	85	8	ADH26390	Novel hum	755	42	87.5	281	3	AAy87579	Human Fas
683	45	93.8	85	8	ADH39018	Novel hum	756	42	87.5	281	3	AAy87574	Human Fas
684	45	93.8	85	8	ADG63586	Human sec	757	42	87.5	281	3	AAy87582	Human Fas
685	45	93.8	85	8	ADH33359	Human PRO	758	42	87.5	281	3	AAy87578	Human Fas
686	45	93.8	85	8	ADH43669	Human PRO	759	42	87.5	281	3	AAy87581	Human Fas
687	45	93.8	85	8	ADG34108	Novel hum	760	42	87.5	281	3	AAy87581	Human Fas
688	45	93.8	85	8	ADH33578	Human PRO	761	42	87.5	281	3	AAy87581	Human Fas
689	45	93.8	85	8	ADH69672	Human PRO	762	42	87.5	281	3	AAy87581	Human Fas
690	45	93.8	85	8	ADH29833	Novel hum	763	42	87.5	281	3	AAy87581	Human Fas
691	45	93.8	85	8	ADH29833	Novel hum	764	42	87.5	281	3	AAy87581	Human Fas
692	45	93.8	85	8	ADH27230	Novel hum	765	42	87.5	281	3	AAy87581	Human Fas
693	45	93.8	85	8	ADH30114	Human PRO	766	42	87.5	281	3	AAy87581	Human Fas
694	45	93.8	85	8	ADH64869	Human PRO	767	42	87.5	281	3	AAy87581	Human Fas
695	45	93.8	85	8	ADH66588	Human PRO	768	42	87.5	281	3	AAy87581	Human Fas
696	45	93.8	85	8	ADH31765	Novel hum	769	42	87.5	281	3	AAy87581	Human Fas
697	45	93.8	85	8	ADH36812	Novel hum	770	42	87.5	281	3	AAy87581	Human Fas
698	45	93.8	85	8	ADH40617	Novel hum	771	42	87.5	281	3	AAy87581	Human Fas
699	45	93.8	85	8	ADH38225	Novel hum	772	42	87.5	281	3	AAy87581	Human Fas
700	45	93.8	85	8	ADP19523	Human sec	773	42	87.5	281	3	AAy87581	Human Fas
701	45	93.8	85	8	ADP19346	Human sec	774	42	87.5	281	3	AAy87581	Human Fas
702	45	93.8	85	8	ADP19345	Human sec	775	42	87.5	281	3	AAy87581	Human Fas
703	45	93.8	85	8	ADU73028	cDNA 48-1	776	42	87.5	281	3	AAy87581	Human Fas
704	45	93.8	85	9	ADZ99302	Secreted	777	42	87.5	281	3	AAy87581	Human Fas
705	45	93.8	86	2	AAy36257	Human sec	778	42	87.5	281	3	AAy87581	Human Fas
706	45	93.8	86	6	ADH11687	Human nov	779	42	87.5	281	3	AAy87581	Human Fas
707	45	93.8	122	6	ADH12081	Human nov	780	42	87.5	281	3	AAy87581	Human Fas
708	44	91.7	12	2	AAy93355	FYN prote	781	42	87.5	281	3	AAy87581	Human Fas
709	44	91.7	112	8	ABM80911	Tumour-as	782	42	87.5	281	3	AAy87581	Human Fas
710	44	91.7	147	8	ADH09417	Human PRO	783	42	87.5	281	3	AAy87581	Human Fas
711	43	89.6	12	2	AAy93347	SRC prote	784	42	87.5	281	3	AAy87581	Human Fas
712	43	89.6	209	7	ADH62971	Human apo	785	42	87.5	281	3	AAy87581	Human Fas
713	43	89.6	209	7	ADH62971	Human apo	786	42	87.5	281	3	AAy87581	Human Fas
714	43	89.6	318	4	ABG19764	Novel hum	787	42	87.5	281	3	AAy87581	Human Fas
715	43	89.6	462	3	ABO7572	Protein e	788	42	87.5	281	3	AAy87581	Human Fas
716	43	89.6	514	8	ADN05033	Antipor	789	42	87.5	281	3	AAy87581	Human Fas
717	43	89.6	514	8	ADP25283	PRO poly	790	42	87.5	281	3	AAy87581	Human Fas
718	43	89.6	1268	4	ABG7095	Drosophil	791	42	87.5	281	3	AAy87581	Human Fas
719	42	87.5	19	2	AAW05473	SH3-bind	792	42	87.5	281	3	AAy87581	Human Fas
720	42	87.5	127	5	AAE29154	Human non	793	42	87.5	281	3	AAy87581	Human Fas
721	42	87.5	158	7	ABO76210	Pseudomon	794	42	87.5	281	3	AAy87581	Human Fas
722	42	87.5	170	7	ABO72623	Pseudomon	795	42	87.5	281	3	AAy87581	Human Fas
723	42	87.5	173	8	ADY22951	Plant ful	796	42	87.5	281	3	AAy87581	Human Fas
724	42	87.5	210	8	ADH87621	Plant ful	797	42	87.5	281	3	AAy87581	Human Fas
725	42	87.5	213	8	ADH88446	Plant ful	798	42	87.5	281	3	AAy87581	Human Fas
726	42	87.5	258	2	AAy04371	Human Fas	799	42	87.5	281	3	AAy87581	Human Fas
727	42	87.5	258	2	ADG14424	Human Fas	800	42	87.5	281	3	AAy87581	Human Fas
728	42	87.5	261	2	AAy28597	Fas lig	801	42	87.5	281	3	AAy87581	Human Fas
729	42	87.5	265	2	AAW48954	Non-cleav	802	42	87.5	281	3	AAy87581	Human Fas
730	42	87.5	267	8	ADJ63977	Human Fas	803	42	87.5	281	3	AAy87581	Human Fas
731	42	87.5	267	8	ADL71817	Human Fas	804	42	87.5	281	3	AAy87581	Human Fas
732	42	87.5	267	8	ADN07588	Human Fas	805	42	87.5	281	3	AAy87581	Human Fas
733	42	87.5	268	2	AAW48953	Non-cleav	806	42	87.5	281	3	AAy87581	Human Fas
734	42	87.5	271	2	AAy28596	Fas lig	807	42	87.5	281	3	AAy87581	Human Fas
735	42	87.5	277	2	AAy04372	Human Fas	808	42	87.5	281	3	AAy87581	Human Fas
736	42	87.5	277	2	AAy28595	Fas lig	809	42	87.5	281	3	AAy87581	Human Fas
737	42	87.5	278	6	ABR82203	Chimeric	810	42	87.5	281	3	AAy87581	Human Fas
738	42	87.5	280	8	ADJ36211	Self-coal	811	42	87.5	281	3	AAy87581	Human Fas
739	42	87.5	281	2	AAy72817	Human Fas	812	42	87.5	281	3	AAy87581	Human Fas
740	42	87.5	281	2	AAy79097	Human Fas	813	42	87.5	281	3	AAy87581	Human Fas
741	42	87.5	281	2	AAH88356	Human Fas	814	42	87.5	281	3	AAy87581	Human Fas
742	42	87.5	281	2	AAH38104	Human Fas	815	42	87.5	281	3	AAy87581	Human Fas
743	42	87.5	281	2	AAW27143	Human Fas	816	42	87.5	281	3	AAy87581	Human Fas
744	42	87.5	281	2	AAW75959	Human Fas	817	42	87.5	281	3	AAy87581	Human Fas
745	42	87.5	281	2	AAW49105	Fas Lig	818	42	87.5	281	3	AAy87581	Human Fas
746	42	87.5	281	2	AAy98071	Human Fas	819	42	87.5	281	3	AAy87581	Human Fas
747	42	87.5	281	2	AAy04373	Human Fas	820	42	87.5	281	3	AAy87581	Human Fas
748	42	87.5	281	2	AAW95041	Human Fas	821	42	87.5	281	3	AAy87581	Human Fas
749	42	87.5	281	2	AAy28594	Wild type	822	42	87.5	281	3	AAy87581	Human Fas
750	42	87.5	281	3	AAy87580	Human Fas	823	42	87.5	281	3	AAy87581	Human Fas
751	42	87.5	281	3	AAy87577	Human Fas	824	42	87.5	281	3	AAy87581	Human Fas
752	42	87.5	281	3	AAy87569	Human Fas	825	42	87.5	281	3	AAy87581	Human Fas
753	42	87.5	281	3	AAy87575	Human Fas	826	42	87.5	281	3	AAy87581	Human Fas
754	42	87.5	281	3	AAy87576	Human Fas	827	42	87.5	281	3	AAy87581	Human Fas

828	41	85.4	12	2	AAR93360	Aar93360	LYN prote	901	41	85.4	230	8	ADY06849	Plant ful
829	41	85.4	12	2	AAR93363	Aar93363	LYN prote	902	41	85.4	278	5	ABY78672	C. elegans
830	41	85.4	12	2	AAR93343	Aar93343	SRC prote	903	41	85.4	288	8	ADY79901	Plant ful
831	41	85.4	12	2	AAR93348	Aar93348	SRC prote	904	41	85.4	292	4	ABY61267	Mature mo
832	41	85.4	12	3	AAB17253	Aab17253	SH3 antag	905	41	85.4	292	6	ABY11233	Mature mo
833	41	85.4	12	4	AAB50782	Aab50782	Human CAM	906	41	85.4	296	8	ADY45297	Rice isop
834	41	85.4	12	5	ABY73246	Abb73246	Src homol	907	41	85.4	303	5	ABY78673	C. elegans
835	41	85.4	12	6	ADA08218	Ada08218	Human SRC	908	41	85.4	313	4	AAB61265	Mouse TAN
836	41	85.4	12	7	ADJ73400	Adj73400	SH3 antag	909	41	85.4	313	4	AAB61278	Mouse TAN
837	41	85.4	12	8	ADJ53034	Adj53034	CHI delet	910	41	85.4	313	4	AAB61280	Mouse TAN
838	41	85.4	12	8	ADJ51995	Adj51995	CHI delet	911	41	85.4	313	4	AAB61277	Mouse TAN
839	41	85.4	13	2	AAB05483	Aab05483	SH3-bind	912	41	85.4	313	4	AAB61279	Mouse TAN
840	41	85.4	14	2	AAB11099	Aab11099	Src SH3 d	913	41	85.4	313	6	ABY11245	Glycoprot
841	41	85.4	14	2	AAB25517	Aab25517	SH3 synth	914	41	85.4	313	6	ABY11243	Glycoprot
842	41	85.4	14	3	AAB08530	Aab08530	Amino aci	915	41	85.4	313	6	ABY11231	Glycoprot
843	41	85.4	14	4	AAB36691	Aab36691	ProS pep	916	41	85.4	313	6	ABY11231	Glycoprot
844	41	85.4	14	5	AAG66171	Aag66171	c-Src pep	917	41	85.4	313	4	ABY11244	Glycoprot
845	41	85.4	15	2	AAR93470	Aar93470	GST-SRC p	918	41	85.4	319	4	AAB61257	Mature hu
846	41	85.4	15	2	AAR93463	Aar93463	GST-P13K	919	41	85.4	319	4	AAB49403	Human gly
847	41	85.4	15	2	AAR93465	Aar93465	GST-P13K	920	41	85.4	319	6	ABY11223	Human TAN
848	41	85.4	15	2	AAR93464	Aar93464	GST-P13K	921	41	85.4	319	8	ADP08399	Human gly
849	41	85.4	15	2	AAR93462	Aar93462	GST-P13K	922	41	85.4	321	8	ADQ39541	Human myo
850	41	85.4	15	2	AAR93466	Aar93466	GST-P13K	923	41	85.4	325	5	ABY78671	C. elegans
851	41	85.4	15	2	AAM05069	Aam05069	Proline-r	924	41	85.4	339	4	AAB61273	Human TAN
852	41	85.4	15	2	AAM05066	Aam05066	Proline-r	925	41	85.4	339	4	AAB61255	Human TAN
853	41	85.4	15	2	AAM05070	Aam05070	Proline-r	926	41	85.4	339	4	AAB61274	Human TAN
854	41	85.4	15	2	AAM38925	Aam38925	Peptide r	927	41	85.4	339	4	AAB61275	Human TAN
855	41	85.4	15	2	AAM38966	Aam38966	Peptide r	928	41	85.4	339	4	AAB61276	Human TAN
856	41	85.4	15	2	AAM38958	Aam38958	Peptide r	929	41	85.4	339	4	AAY72790	Human pla
857	41	85.4	15	3	AAB20793	Aab20793	Src, Fyn,	930	41	85.4	339	4	AAB31668	Amino aci
858	41	85.4	16	2	AAM25448	Aam25448	Grb2 N-te	931	41	85.4	339	5	ABY78674	C. elegans
859	41	85.4	21	2	AAM05452	Aam05452	SH3-bind	932	41	85.4	339	6	ADA84105	Human GP6
860	41	85.4	22	2	AAM37660	Aam37660	PPPYP mot	933	41	85.4	339	6	ABY11221	Human TAN
861	41	85.4	22	7	ADB49286	Abd49286	Novel WW	934	41	85.4	339	6	ABY11239	Glycoprot
862	41	85.4	25	9	ABY07370	Aeb07370	Signal tr	935	41	85.4	339	6	ABY11240	Glycoprot
863	41	85.4	27	4	AAB61270	Aab61270	Mouse TAN	936	41	85.4	339	6	ABY11241	Glycoprot
864	41	85.4	27	6	ABY11236	Abu11236	Mouse TAN	937	41	85.4	339	6	ABY11242	Glycoprot
865	41	85.4	30	9	ABY07371	Aeb07371	Signal tr	938	41	85.4	339	8	ADH22420	Human pro
866	41	85.4	40	4	AAM18998	Aam18998	Peptide #	939	41	85.4	386	5	ABY78660	C. elegans
867	41	85.4	40	4	ABY38188	Abb38188	Peptide #	940	41	85.4	386	5	ABY78667	C. elegans
868	41	85.4	40	4	AAM31617	Aab31617	Peptide #	941	41	85.4	386	5	ABY78668	C. elegans
869	41	85.4	40	4	ABY23379	Abb23379	Protein #	942	41	85.4	386	5	ABY78665	C. elegans
870	41	85.4	40	4	AAM71338	Aam71338	Human bon	943	41	85.4	386	5	ABY78675	C. elegans
871	41	85.4	40	4	AAM58822	Aam58822	Human bra	944	41	85.4	386	5	ABY78665	C. elegans
872	41	85.4	40	4	ABY3046	Abg53046	Human liv	945	41	85.4	386	5	ABY78667	C. elegans
873	41	85.4	40	5	ABY41141	Abg41141	Human pep	946	41	85.4	386	5	ABY78664	C. elegans
874	41	85.4	51	6	ABY11228	Abu11228	Human TAN	947	41	85.4	386	5	ABY78661	C. elegans
875	41	85.4	51	6	ABY11228	Abu11228	Human TAN	948	41	85.4	386	5	ABY78661	C. elegans
876	41	85.4	73	5	ABY09414	Abb09414	Ovarian c	949	41	85.4	398	9	ABY20718	TRP domai
877	41	85.4	73	5	ABY08093	Aao18093	Human ova	950	41	85.4	398	9	ABY07376	Signal tr
878	41	85.4	73	7	ADY10871	Adm10871	O573S_411	951	41	85.4	401	4	AAM79810	Human pro
879	41	85.4	73	8	ADY11201	Adm11201	Human ova	952	41	85.4	432	8	ADY87392	Plant ful
880	41	85.4	73	8	ADY43462	Adm43462	Human ova	953	41	85.4	459	7	ADY63349	Rat Prote
881	41	85.4	78	5	ABY78662	Abb78662	C. elegans	954	41	85.4	461	8	ADY60064	CRH signa
882	41	85.4	83	3	ABY0669	Abb0669	Human ORF	955	41	85.4	462	3	ABY18668	A human r
883	41	85.4	83	5	ABY06901	Abp06901	Human ORF	956	41	85.4	462	4	AAM78826	Human pro
884	41	85.4	88	8	ADY66827	Adx66827	Plant ful	957	41	85.4	462	5	ABY70175	Human pre
885	41	85.4	108	4	ABY36682	Abb36682	Mammalian	958	41	85.4	462	6	ABY57649	Different
886	41	85.4	112	5	ABY78663	Abb78663	C. elegans	959	41	85.4	462	7	ADY63351	Human pro
887	41	85.4	127	4	ABY01550	Aao1550	Human pol	960	41	85.4	462	8	ADY12698	Human ste
888	41	85.4	146	5	ABY78666	Abb78666	C. elegans	961	41	85.4	480	4	ABY36684	Mammalian
889	41	85.4	152	8	ADY22975	Ady22975	Plant ful	962	41	85.4	495	6	ABY32076	Human cer
890	41	85.4	156	2	AAR13587	Aar13587	PLRV vira	963	41	85.4	495	7	ADY34349	Human hyp
891	41	85.4	161	4	ABY36683	Abb36683	Mammalian	964	41	85.4	495	8	ADY24979	PRO poly
892	41	85.4	162	5	ABY78669	Abb78669	C. elegans	965	41	85.4	495	8	ADY06363	Novel bro
893	41	85.4	176	3	ABY50994	Aay50994	Potato pr	966	41	85.4	513	6	ABY96232	Human nuc
894	41	85.4	183	8	ADY90517	Adx90517	Plant ful	967	41	85.4	535	4	ABY66593	Drosophil
895	41	85.4	186	8	ADY48382	Adx48382	Streptoco	968	41	85.4	548	4	ABY5289	Drosophil
896	41	85.4	186	8	ADY90922	Adx90922	Plant ful	969	41	85.4	562	5	ABY69088	Human pol
897	41	85.4	203	4	ABY15907	Abg15907	Novel hum	970	41	85.4	710	4	AAM43571	Human pol
898	41	85.4	206	4	ABY13470	Abg13470	Novel hum	971	41	85.4	710	8	ADY24592	Human PRO
899	41	85.4	219	8	ADY79537	Ady79537	Plant ful	972	41	85.4	717	4	AAM18383	Human end
900	41	85.4	225	4	AAB36690	Abb36690	Mammalian	973	41	85.4	760	4	AAM93625	Human pol

974 41 85.4 760 8 ADL31429 Human pro
 975 41 85.4 782 8 ADT56202 Plant pol
 976 41 85.4 847 5 ABG34076 Human pro
 977 41 85.4 847 6 ADA01362 Human pro
 978 41 85.4 847 6 ADA3791 Human sec
 979 41 85.4 847 6 ADA3559 Human sec
 980 41 85.4 847 6 ADA01234 Human pro
 981 41 85.4 847 7 ADA01118 Human sec
 982 41 85.4 847 7 ADA43675 Human sec
 983 41 85.4 847 7 ADA06937 Human pro
 984 41 85.4 847 7 ADA08425 Human pro
 985 41 85.4 847 7 ADB99718 Human pro
 986 41 85.4 847 7 ADB87001 Human pro
 987 41 85.4 847 7 ADB66156 Human sec
 988 41 85.4 847 7 ADB99834 Human pro
 989 41 85.4 847 7 ADB99489 Novel hum
 990 41 85.4 847 7 ADB66040 Human sec
 991 41 85.4 847 7 ADC23438 Human tra
 992 41 85.4 847 7 ADC26131 Human pro
 993 41 85.4 847 7 ADE04958 Human pro
 994 41 85.4 847 7 ADE11264 Human pro
 995 41 85.4 847 7 ADE88195 Human pro
 996 41 85.4 847 7 ADE95490 Human sec
 997 41 85.4 847 7 ADE06420 Human pro
 998 41 85.4 847 7 ADE38195 Human pro
 999 41 85.4 847 7 ADD88311 Human pro
 1000 41 85.4 847 7 ADD90892 Human sec

ALIGNMENTS

RESULT 1
 AAR933346
 ID AAR93346 standard; peptide; 12 AA.

XX AC AAR933346;
 XX AC AAR933346;
 DT 24-APR-1996 (first entry)
 XX SRC/LYN protein tyrosine kinase derived peptide.
 XX SH3 ligand; SH3 binding agent; biased phage library;
 KW recognition sequence; src SH3 domain; Paget's disease; restenosis;
 KW rheumatoid arthritis; gout; neutrophil oxidase; tyrosine kinase; p47;
 KW p67; complex; chronic myelogenous leukaemia; cancer.
 XX Synthetic.
 OS Synthetic.
 XX WO9524419-A1.
 XX WO9524419-A1.
 XX 14-SEP-1995.
 XX 13-MAR-1995; 95WO-US003208.
 XX 11-MAR-1994; 94US-00209835.
 XX 06-JAN-1995; 95US-00369832.
 XX (ARIA-) ARIAD PHARM INC.
 XX Rickles RJ, Brugge JS, Botfield MC, Zoller MJ;
 WPI; 1995-328231/42.

XX Identification of peptide(s) binding specifically to SH3 domains - for
 PT use in inhibiting interactions mediated by SH3 domains in treatment of
 PT e.g. osteoporosis and cancer.
 XX Disclosure; Fig 2; 74pp; English.

XX The sequences given in AAR93343-68 represent peptides which are SH3
 CC ligands/SH3 binding agents. They represent a biased phage library which
 CC comprises six random amino acids flanking the hexapeptide RSLRPL- which

CC was identified as a recognition sequence for the src SH3 domain. These
 CC sequences were identified using the method of the invention. The method
 CC comprises contacting the SH3 domain with a mixture of peptides under
 CC conditions permitting a ligand to bind to an SH3 domain to form a
 CC complex. Any unbound peptides are removed and the complexed peptide
 CC ligands are dissociated from the complexes. The selected peptides are
 CC enriched by re-contacting them with the SH3 domain and then candidates
 CC which bind to the SH3 domain are detected. The isolated SH3 binding
 CC peptides may be used in the diagnosis, prevention and treatment of
 CC conditions or diseases resulting from cellular processes mediated by an
 CC SH3-based interaction. Such diseases include Paget's disease. Other
 CC conditions treatable with these peptides include restenosis, rheumatoid
 CC arthritis, gout and other problems in which an SH3 of neutrophil oxidase
 CC p47 and p67 complex is implicated, etc
 XX Sequence 12 AA;

Query Match 100.0%; Score 48; DB 2; Length 12;
 Best Local Similarity 88.9%; Pred. No. 6.3;
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 RPLPPLXP 12
 |||||
 Db 4 RPLPPLSP 12

RESULT 2
 AAR933356
 ID AAR93356 standard; peptide; 12 AA.
 XX AC AAR93356;
 XX AC AAR93356;
 DT 24-APR-1996 (first entry)
 XX FYN/LYN protein tyrosine kinase derived peptide.

XX SH3 ligand; SH3 binding agent; biased phage library;
 KW recognition sequence; src SH3 domain; Paget's disease; restenosis;
 KW rheumatoid arthritis; gout; neutrophil oxidase; tyrosine kinase; p47;
 KW p67; complex; chronic myelogenous leukaemia; cancer.
 XX Synthetic.

XX WO9524419-A1.
 XX 14-SEP-1995.
 XX 13-MAR-1995; 95WO-US003208.
 XX 11-MAR-1994; 94US-00209835.
 XX 06-JAN-1995; 95US-00369832.
 XX (ARIA-) ARIAD PHARM INC.

XX Rickles RJ, Brugge JS, Botfield MC, Zoller MJ;
 WPI; 1995-328231/42.
 XX Identification of peptide(s) binding specifically to SH3 domains - for
 PT use in inhibiting interactions mediated by SH3 domains in treatment of
 PT e.g. osteoporosis and cancer.
 XX Disclosure; Fig 2; 74pp; English.

XX The sequences given in AAR93343-68 represent peptides which are SH3
 CC ligands/SH3 binding agents. They represent a biased phage library which
 CC comprises six random amino acids flanking the hexapeptide RSLRPL- which
 CC was identified as a recognition sequence for the src SH3 domain. These
 CC sequences were identified using the method of the invention. The method
 CC comprises contacting the SH3 domain with a mixture of peptides under
 CC conditions permitting a ligand to bind to an SH3 domain to form a
 CC complex. Any unbound peptides are removed and the complexed peptide
 CC ligands are dissociated from the complexes. The selected peptides are

CC enriched by re-contacting them with the SH3 domain and then candidates
 CC which bind to the SH3 domain are detected. The isolated SH3 binding
 CC peptides may be used in the diagnosis, prevention and treatment of
 CC conditions or diseases resulting from cellular processes mediated by an
 CC SH3-based interaction. Such diseases include Paget's disease. Other
 CC conditions treatable with these peptides include restenosis, rheumatoid
 CC arthritis, gout and other problems in which an SH3 of neutrophil oxidase
 CC p47 and p67 complex is implicated, etc

XX Sequence 12 AA;

Query Match 100.0%; Score 48; DB 2; Length 12;
 Best Local Similarity 88.9%; Pred. No. 6.3;
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 RPLPPLPXP 12
 |||||
 DB 4 RPLPPLPAP 12

RESULT 3
 AAR93471
 ID AAR93471 standard; peptide; 15 AA.

XX
 AC AAR93471;
 DT 09-MAY-1996 (first entry)

DE GST-SRC protein tyrosine kinase derived peptide #5.

XX SH3 ligand; SH3 binding agent; biased phage library;
 KW recognition sequence; src SH3 domain; Paget's disease; restenosis;
 KW rheumatoid arthritis; gout; neutrophil oxidase; tyrosine kinase; p47;
 KW p67; complex; chronic myelogenous leukaemia; cancer.

XX Synthetic.

XX WO9524419-A1.

XX 14-SEP-1995.

XX 13-MAR-1995; 95WO-US003208.

PR 11-MAR-1994; 94US-00209835.

PR 06-JAN-1995; 95US-00369832.

XX (ARIA-) ARIAD PHARM INC.

XX Rickles RJ, Brugge JS, Botfield MC, Zoller MJ;

XX WPI; 1995-328231/42.

XX Identification of peptide(s) binding specifically to SH3 domains - for
 PT use in inhibiting interactions mediated by SH3 domains in treatment of
 PT e.g. osteoporosis and cancer.

XX Disclosure; Fig 5; 74pp; English.

XX The sequences given in AAR93457-71 represent peptides which are SH3
 CC ligands/SH3 binding agents. They represent a biased phage library which
 CC comprises five random amino acids flanking the decapeptide -RSRLPPLP or
 CC derivatives of this, which was identified as a recognition sequence for
 CC the src SH3 domain. These sequences were identified using the method of
 CC the invention. The method comprises contacting the SH3 domain with a
 CC mixture of peptides under conditions permitting a ligand to bind to an
 CC SH3 domain to form a complex. Any unbound peptides are removed and the
 CC complexed peptide ligands are dissociated from the complexes. The
 CC selected peptides are enriched by re-contacting them with the SH3 domain
 CC and then candidates which bind to the SH3 domain are detected. The
 CC isolated SH3 binding peptides may be used in the diagnosis, prevention
 CC and treatment of conditions or diseases resulting from cellular processes
 CC mediated by an SH3-based interaction. Such diseases include Paget's
 CC disease. Other conditions treatable with these peptides include

CC restenosis, rheumatoid arthritis, gout and other problems in which an SH3
 CC of neutrophil oxidase p47 and p67 complex is implicated, etc

XX Sequence 15 AA;

Query Match 100.0%; Score 48; DB 2; Length 15;
 Best Local Similarity 88.9%; Pred. No. 7.7;
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 RPLPPLPXP 12
 |||||
 DB 4 RPLPPLPFP 12

RESULT 4
 AAB21130
 ID AAB21130 standard; peptide; 11 AA.

XX AC AAB21130;

XX 19-JAN-2001 (first entry)

XX Src homology 3 domain binding peptide #7.

XX Src homology domain 3; SH3; protein-protein interaction; cancer;
 KW signal transduction inhibition; immune suppression-associated disease.

XX Synthetic.

XX WO2000047607-A1.

XX 17-AUG-2000.

XX 12-FEB-2000; 2000WO-KR000107.

XX 12-FEB-1999; 99AU-00008643.

XX 02-JUN-1999; 99KR-00020282.

XX (YOON/) YOON J H.

XX (HANY/) HAN Y T.

XX Yoon JH, Han YT, Lee KY;

XX WPI; 2000-533010/48.

XX Synthetic peptides useful for treating cancers and immunosuppressive
 PT disorders by disrupting interactions of the SH (src homology) 2 and SH3
 PT motifs of Src family kinase proteins.

XX Claim 3; Page 34; 40pp; English.

XX The present sequence is a synthetic peptide which has a high affinity for
 CC the src homology 3 (SH3) domain of protein kinases. Protein kinases are
 CC involved in signal transduction pathways, and this peptide can be used to
 CC inhibit these, by disrupting protein-protein interactions, in the
 CC treatment of cancer, particularly hepatocellular carcinoma, cervical
 CC cancer, colon adenocarcinoma and fibrosarcoma, and immune suppression-
 CC associated diseases

XX Sequence 11 AA;

Query Match 97.9%; Score 47; DB 3; Length 11;
 Best Local Similarity 88.9%; Pred. No. 7.9;
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 RPLPPLPXP 12
 |||||
 DB 2 RPLPPLPLP 10

RESULT 5
 AAB21125
 ID AAB21125 standard; peptide; 11 AA.

XX AAB21125;
 XX AC
 XX 19-JAN-2001 (first entry)
 XX DE
 XX Src homology 3 domain binding peptide #2.
 XX KW
 XX Src homology domain 3; SH3; protein-protein interaction; cancer;
 XX KW signal transduction inhibition; immune suppression-associated disease.
 XX OS
 XX Synthetic.
 XX PN W0200047607-A1.
 XX XX
 XX 17-AUG-2000.
 XX XX
 XX 12-FEB-2000; 2000WO-KR000107.
 XX XX
 XX 12-FEB-1999; 99AU-00008643.
 XX PR 02-JUN-1999; 99KR-00020282.
 XX XX
 XX (YOON/) YOON J H.
 XX PA (HANY/) HAN Y T.
 XX XX
 XX Yoon JH, Han YT, Lee KY;
 XX PI WPI; 2000-533010/48.
 XX DR
 XX Synthetic peptides useful for treating cancers and immunosuppressive
 XX PT disorders by disrupting interactions of the SH (Src homology) 2 and SH3
 XX PT motifs of Src family kinase proteins.
 XX XX
 XX Claim 3; Page 32; 40pp; English.
 XX XX
 XX The present sequence is a synthetic peptide which has a high affinity for
 XX CC the src homology 3 (SH3) domain of protein kinases. Protein kinases are
 XX CC involved in signal transduction pathways, and this peptide can be used to
 XX CC inhibit these, by disrupting protein-protein interactions, in the
 XX CC treatment of cancer, particularly hepatocellular carcinoma, cervical
 XX CC cancer, colon adenocarcinoma and fibrosarcoma, and immune suppression-
 XX CC associated diseases
 XX XX
 XX Sequence 11 AA;
 Query Match 97.9%; Score 47; DB 3; Length 11;
 Best Local Similarity 88.9%; Pred. No. 7.9;
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 4 RPLPPLPXP 12
 Db 2 RPLPPLPNP 10
 RESULT 6
 AAR93351
 ID AAR93351 standard; peptide; 12 AA.
 XX AC
 XX AAR93351;
 XX XX
 XX 24-APR-1996 (first entry)
 XX DT
 XX FYN protein tyrosine kinase derived peptide #3.
 XX DE
 XX SH3 ligand; SH3 binding agent; biased phage library;
 XX KW recognition sequence; src SH3 domain; Paget's disease; restenosis;
 XX KW rheumatoid arthritis; gout; neutrophil oxidase; tyrosine kinase; p47;
 XX KW p67; complex; chronic myelogenous leukaemia; cancer.
 XX XX
 XX Synthetic.
 XX OS
 XX W09524419-A1.
 XX PN
 XX 14-SEP-1995.
 XX PD

XX 13-MAR-1995; 95WO-US003208.
 XX PF
 XX 11-MAR-1994; 94US-00209835.
 XX PR
 XX 06-JAN-1995; 95US-00369832.
 XX XX
 XX (ARIA-) ARIAD PHARM INC.
 XX PA
 XX Rickles RJ, Brugge JS, Botfield MC, Zoller MJ;
 XX PI WPI; 1995-328231/42.
 XX DR
 XX Identification of peptide(s) binding specifically to SH3 domains - for
 XX PT use in inhibiting interactions mediated by SH3 domains in treatment of
 XX PT e.g. osteoporosis and cancer.
 XX XX
 XX Disclosure; Fig 2; 74pp; English.
 XX PS
 XX The sequences given in AAR93343-68 represent peptides which are SH3
 XX CC ligands/SH3 binding agents. They represent a biased phage library which
 XX CC comprises six random amino acids flanking the hexapeptide RSLAPL- which
 XX CC was identified as a recognition sequence for the src SH3 domain. These
 XX CC sequences were identified using the method of the invention. The method
 XX CC comprises contacting the SH3 domain with a mixture of peptides under
 XX CC conditions permitting a ligand to bind to an SH3 domain to form a
 XX CC complex. Any unbound peptides are removed and the complexed peptide
 XX CC ligands are dissociated from the complexes. The selected peptides are
 XX CC enriched by re-contacting them with the SH3 domain and then candidates
 XX CC which bind to the SH3 domain are detected. The isolated SH3 binding
 XX CC peptides may be used in the diagnosis, prevention and treatment of
 XX CC conditions or diseases resulting from cellular processes mediated by an
 XX CC SH3-based interaction. Such diseases include Paget's disease. Other
 XX CC conditions treatable with these peptides include restenosis, rheumatoid
 XX CC arthritis, gout and other problems in which an SH3 of neutrophil oxidase
 XX CC p47 and p67 complex is implicated, etc
 XX XX
 XX Sequence 12 AA;
 Query Match 97.9%; Score 47; DB 2; Length 12;
 Best Local Similarity 88.9%; Pred. No. 8.5;
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 4 RPLPPLPXP 12
 Db 4 RPLPPLPVP 12
 RESULT 7
 AAR93359
 ID AAR93359 standard; peptide; 12 AA.
 XX XX
 XX AAR93359;
 XX AC
 XX 24-APR-1996 (first entry)
 XX DT
 XX LYN protein tyrosine kinase derived peptide #1.
 XX DE
 XX SH3 ligand; SH3 binding agent; biased phage library;
 XX KW recognition sequence; src SH3 domain; Paget's disease; restenosis;
 XX KW rheumatoid arthritis; gout; neutrophil oxidase; tyrosine kinase; p47;
 XX KW p67; complex; chronic myelogenous leukaemia; cancer.
 XX XX
 XX Synthetic.
 XX OS
 XX W09524419-A1.
 XX PN
 XX 14-SEP-1995.
 XX PD
 XX 13-MAR-1995; 95WO-US003208.
 XX PF
 XX 11-MAR-1994; 94US-00209835.
 XX PR
 XX 06-JAN-1995; 95US-00369832.
 XX XX

PA (ARIA-) ARIAD PHARM INC.
 XX Rickles RJ, Brugge JS, Botfield MC, Zoller MJ;
 PI WPI; 1995-328231/42.
 XX
 XX Identification of peptide(s) binding specifically to SH3 domains - for
 PT use in inhibiting interactions mediated by SH3 domains in treatment of
 PT e.g. osteoporosis and cancer.
 XX
 XX Disclosure; Fig 2; 74pp; English.
 XX
 XX The sequences given in AAR93343-68 represent peptides which are SH3
 CC ligands/SH3 binding agents. They represent a biased phage library which
 CC comprises six random amino acids flanking the hexapeptide RSLRPL- which
 CC was identified as a recognition sequence for the src SH3 domain. These
 CC sequences were identified using the method of the invention. The method
 CC comprises contacting the SH3 domain with a mixture of peptides under
 CC conditions permitting a ligand to bind to an SH3 domain to form a
 CC complex. Any unbound peptides are removed and the complexed peptide
 CC ligands are dissociated from the complexes. The selected peptides are
 CC enriched by re-contacting them with the SH3 domain and then candidates
 CC which bind to the SH3 domain are detected. The isolated SH3 binding
 CC peptides may be used in the diagnosis, prevention and treatment of
 CC conditions or diseases resulting from cellular processes mediated by an
 CC SH3-based interaction. Such diseases include Paget's disease. Other
 CC conditions treatable with these peptides include restenosis, rheumatoid
 CC arthritis, gout and other problems in which an SH3 of neutrophil oxidase
 CC p47 and p67 complex is implicated, etc
 XX
 XX Sequence 12 AA;
 SQ

Query Match 97.9%; Score 47; DB 2; Length 12;
 Best Local Similarity 88.9%; Pred. No. 8.5;
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 RPLPPLPXP 12
 |||||
 Db 4 RPLPPLPXP 12

RESULT 9
 AAB17254
 ID AAB17254 standard; peptide; 12 AA.
 XX
 AC AAB17254;
 XX
 DT 31-OCT-2000 (first entry)
 XX
 DE SH3 antagonist peptide sequence SEQ ID NO:310.
 XX
 KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
 KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
 KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
 KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;
 KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
 KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
 KW thrombosis; pharmaceutical.
 XX
 OS Synthetic.
 XX
 PN WO200024782-A2.
 XX
 PD 04-MAY-2000.
 XX
 PF 25-OCT-1999; 99WO-US025044.
 XX
 PR 23-OCT-1998; 98US-0105371P.
 PR 22-OCT-1999; 99US-00428082.
 XX
 XX (AMGE-) AMGEN INC.
 PA
 XX Feige U, Liu C, Cheetham J, Boone TC;
 PI WPI; 2000-350702/30.
 XX
 XX Novel composition of matter comprising an Fc domain and pharmacologically
 PT active peptides, useful for treating cancer and autoimmune diseases.
 XX

PA (ARIA-) ARIAD PHARM INC.
 XX Rickles RJ, Brugge JS, Botfield MC, Zoller MJ;
 PI WPI; 1995-328231/42.
 XX
 XX Identification of peptide(s) binding specifically to SH3 domains - for
 PT use in inhibiting interactions mediated by SH3 domains in treatment of
 PT e.g. osteoporosis and cancer.
 XX
 XX Disclosure; Fig 2; 74pp; English.
 XX
 XX The sequences given in AAR93343-68 represent peptides which are SH3
 CC ligands/SH3 binding agents. They represent a biased phage library which
 CC comprises six random amino acids flanking the hexapeptide RSLRPL- which
 CC was identified as a recognition sequence for the src SH3 domain. These
 CC sequences were identified using the method of the invention. The method
 CC comprises contacting the SH3 domain with a mixture of peptides under
 CC conditions permitting a ligand to bind to an SH3 domain to form a
 CC complex. Any unbound peptides are removed and the complexed peptide
 CC ligands are dissociated from the complexes. The selected peptides are
 CC enriched by re-contacting them with the SH3 domain and then candidates
 CC which bind to the SH3 domain are detected. The isolated SH3 binding
 CC peptides may be used in the diagnosis, prevention and treatment of
 CC conditions or diseases resulting from cellular processes mediated by an
 CC SH3-based interaction. Such diseases include Paget's disease. Other
 CC conditions treatable with these peptides include restenosis, rheumatoid
 CC arthritis, gout and other problems in which an SH3 of neutrophil oxidase
 CC p47 and p67 complex is implicated, etc
 XX
 XX Sequence 12 AA;
 SQ

Query Match 97.9%; Score 47; DB 2; Length 12;
 Best Local Similarity 88.9%; Pred. No. 8.5;
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 RPLPPLPXP 12
 |||||
 Db 4 RPLPPLPXP 12

RESULT 8
 AAR93344
 ID AAR93344 standard; peptide; 12 AA.
 XX
 AC AAR93344;
 XX
 DT 24-APR-1996 (first entry)
 XX
 DE SRC protein tyrosine kinase derived peptide #2.
 XX
 KW SH3 ligand; SH3 binding agent; biased phage library;
 KW recognition sequence; src SH3 domain; Paget's disease; restenosis;
 KW rheumatoid arthritis; gout; neutrophil oxidase; tyrosine kinase; p47;
 KW p67; complex; chronic myelogenous leukaemia; cancer.
 XX
 OS Synthetic.
 XX
 PN WO9524419-A1.
 XX
 PD 14-SEP-1995.
 XX
 PF 13-MAR-1995; 95WO-US003208.
 XX
 PR 11-MAR-1994; 94US-00209835.
 PR 06-JAN-1995; 95US-00369832.
 XX
 XX (ARIA-) ARIAD PHARM INC.
 PA
 XX Rickles RJ, Brugge JS, Botfield MC, Zoller MJ;
 PI WPI; 1995-328231/42.
 XX

PS Claim 39; Page 304; 608pp; English.

XX The present invention describes composition of matter (I) comprising an
 CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
 CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
 CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
 CC (L2)d-P2-(L3)e-P³, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
 CC P3, and P4 = are each independently sequences of pharmacologically active
 CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,
 CC c, d, e, and f = are each independently 0 or 1, provided that at least 1
 CC of a and b is 1. The composition can have cytostatic, antiasthmatic,
 CC thrombolytic and immunosuppressive activities. DNAs, vectors and host
 CC cells from the present invention can be used for producing pharmaceutical
 CC compositions. The compositions are useful for treating cancer, asthma,
 CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
 CC a Fab domain) can provide a longer half-life or incorporate functions
 CC such as Fc receptor binding, protein A binding, complement fixation, and
 CC possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to
 CC AAB18003 represent nucleotide and amino acid sequences used in the
 CC exemplification of the present invention

XX
 SQ Sequence 12 AA;

Query Match 97.9%; Score 47; DB 3; Length 12;
 Best Local Similarity 100.0%; Pred. No. 8.5;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 RPLPLPLXP 12
 |||||

Db 4 RPLPLPLXP 12
 |||||

RESULT 10

AAB17251

ID AAB17251 standard; peptide; 12 AA.

AC AAB17251;

XX 31-OCT-2000 (first entry)

DE SH3 antagonist peptide sequence SEQ ID NO:307.

XX Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
 KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
 KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
 KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;
 KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
 KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
 KW thrombosis; pharmaceutical.

XX Synthetic.

OS WO200024782-A2.

PN 04-MAY-2000.

PD 25-OCT-1999; 99WO-US025044.

XX 23-OCT-1998; 98US-0105371P.

PR 22-OCT-1999; 99US-00428082.

XX (AMGE-) AMGEN INC.

XX Feige U, Liu C, Cheetham J, Boone TC;

PI WPI; 2000-350702/30.

DR Novel composition of matter comprising an Fc domain and pharmacologically
 XX active peptides, useful for treating cancer and autoimmune diseases.

PT Claim 39; Page 303; 608pp; English.

XX The present invention describes composition of matter (I) comprising an

CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
 CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
 CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
 CC (L2)d-P2-(L3)e-P³, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
 CC P3, and P4 = are each independently sequences of pharmacologically active
 CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,
 CC c, d, e, and f = are each independently 0 or 1, provided that at least 1
 CC of a and b is 1. The composition can have cytostatic, antiasthmatic,
 CC thrombolytic and immunosuppressive activities. DNAs, vectors and host
 CC cells from the present invention can be used for producing pharmaceutical
 CC compositions. The compositions are useful for treating cancer, asthma,
 CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
 CC a Fab domain) can provide a longer half-life or incorporate functions
 CC such as Fc receptor binding, protein A binding, complement fixation, and
 CC possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to
 CC AAB18003 represent nucleotide and amino acid sequences used in the
 CC exemplification of the present invention

XX
 SQ Sequence 12 AA;

Query Match 97.9%; Score 47; DB 3; Length 12;
 Best Local Similarity 100.0%; Pred. No. 8.5;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 RPLPLPLXP 12
 |||||

Db 4 RPLPLPLXP 12
 |||||

RESULT 11

ABB73244

ID ABB73244 standard; peptide; 12 AA.

AC ABB73244;

XX 05-APR-2002 (first entry)

DE Src homology3 (SH3) antagonist peptide SEQ ID NO:307.

XX Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
 KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
 KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;
 KW TPO mimetic peptide; EPO mimetic peptide; BMP; VEGF antagonist;
 KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
 KW cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
 KW antianemic; anorectic; antifertility; haemostatic; dermatological;
 KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
 KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
 KW sleep disorder; neurological degenerative disease; anaemia;
 KW thrombocytopaenia; metastatic tumour; systemic lupus erythematosus;
 KW Fanconi's syndrome.

XX Homo sapiens.

OS Synthetic.

PN WO200183525-A2.

PD 08-NOV-2001.

XX 02-MAY-2001; 2001WO-US014310.

PR 03-MAY-2000; 2000US-00563286.

XX (AMGE-) AMGEN INC.

XX Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;

PI WPI; 2002-130313/17.

DR Novel vehicle-peptide molecule or its multimers useful for treating
 XX inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
 PT diabetic retinopathy, obesity, sleep disorders and infertility.

PS Claim 39; Page 55; 176pp; English.

XX The present invention describes a vehicle-peptide molecule (I) or its
 CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,
 CC cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,
 CC antianaemic, anorectic, antiinfertility, haemostatic, dermatological and
 CC neuroprotective activities. (I) can be used as a therapeutic or
 CC prophylactic agent as well as for screening purposes. (I) is useful for
 CC diagnosing diseases characterised by dysfunction of their associated
 CC protein of interest, for identifying normal or abnormal proteins of
 CC interest, as a part of diagnostic kit to detect the presence of their
 CC proteins of interest in a biological sample. Additionally, (I) is useful
 CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,
 CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders, EPO-
 CC infertility, and neurological degenerative diseases. (I), comprising EPO-
 CC mimetic compounds are useful for treating disorders characterised by low
 CC red blood cell levels such as anaemia. The TPO-mimetic comprising
 CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet
 CC deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic
 CC tumour which result in thrombocytopaenia, systemic lupus erythematosus,
 CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777
 CC represent amino acid and nucleic acid sequences used in the
 CC exemplification of the present invention

XX Sequence 12 AA;

Qy Query Match 97.9%; Score 47; DB 5; Length 12;
 Db Best Local Similarity 100.0%; Pred. No. 8.5;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 RLPPLPXP 12
 Db |||||

RESULT 12

AB73247
 ID ABB73247 standard; peptide; 12 AA.

XX AC ABB73247;

XX DT 05-APR-2002 (first entry)

XX Src homology3 (SH3) antagonist peptide SEQ ID NO:310.

XX Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
 KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
 KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TNP;
 KW TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;
 KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
 KW cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
 KW antianaemic; anorectic; antiinfertility; haemostatic; dermatological;
 KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
 KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
 KW sleep disorder; neurological degenerative disease; anaemia;
 KW thrombocytopaenia; metastatic tumour; systemic lupus erythematosus;
 KW Fanconi's syndrome.

XX Homo sapiens.

OS Synthetic.

XX WO200183525-A2.

XX 08-NOV-2001.

XX 02-MAY-2001; 2001WO-US014310.

XX 03-MAY-2000; 2000US-00563286.

XX (AMGE-) AMGEN INC.

XX Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;

XX WPI; 2002-130313/17.

XX Novel vehicle-peptide molecule or its multimers useful for treating
 PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
 PT diabetic retinopathy, obesity, sleep disorders and infertility.

XX Claim 39; Page 55; 176pp; English.

XX The present invention describes a vehicle-peptide molecule (I) or its
 CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,
 CC cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,
 CC antianaemic, anorectic, antiinfertility, haemostatic, dermatological and
 CC neuroprotective activities. (I) can be used as a therapeutic or
 CC prophylactic agent as well as for screening purposes. (I) is useful for
 CC diagnosing diseases characterised by dysfunction of their associated
 CC protein of interest, for identifying normal or abnormal proteins of their
 CC interest, as a part of diagnostic kit to detect the presence of their
 CC proteins of interest in a biological sample. Additionally, (I) is useful
 CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,
 CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders, EPO-
 CC infertility, and neurological degenerative diseases. (I), comprising EPO-
 CC mimetic compounds are useful for treating disorders characterised by low
 CC red blood cell levels such as anaemia. The TPO-mimetic comprising
 CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet
 CC deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic
 CC tumour which result in thrombocytopaenia, systemic lupus erythematosus,
 CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777
 CC represent amino acid and nucleic acid sequences used in the
 CC exemplification of the present invention

XX Sequence 12 AA;

Qy Query Match 97.9%; Score 47; DB 5; Length 12;
 Db Best Local Similarity 100.0%; Pred. No. 8.5;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 RLPPLPXP 12
 Db |||||

RESULT 13

ADJ73401
 ID ADJ73401 standard; peptide; 12 AA.

XX AC ADJ73401;

XX DT 06-MAY-2004 (first entry)

XX SH3 antagonist peptide sequence SeqID 856.

XX mimetic; CDR mimetibody; gene therapy; transgenic; immune;
 KW cardiovascular; infectious; malignant; neurologic disease; anaemia;
 KW immunomodulator; cardiac; antimicrobial; cytostatic; neuroprotective;
 KW SH3.

XX Synthetic.

XX WO2003084477-A2.

XX 16-OCT-2003.

XX 24-MAR-2003; 2003WO-US009139.

XX 29-MAR-2002; 2002US-0368791P.

XX (CENZ) CENTOCOR INC.

XX Heavner GA, Knight DM, Scallon BJ, Ghayeb J;

XX WPI; 2003-804237/75.

XX New CDR mimetibody comprising a portion of a heavy or light chain
PT variable region comprising human framework or ligand binding region,
PT useful for preparing a composition for treating e.g., immune,
PT cardiovascular or neurologic disease.
XX Disclosure; SEQ ID NO 856; 97pp; English.
PS
XX This invention relates to novel mammalian CDR mimetibodies, specific
CC portions or variants thereof. Specifically, it refers to an antibody
CC fragment where a protein has been inserted into, or replaces a portion
CC of, one or more CDR regions, such that each CDR mimetibody comprises at
CC least one portion of a heavy chain or light chain variable region, which
CC itself comprises at least one human framework region and at least one
CC ligand binding region (LBR). The present invention describes human
CC mimetibodies, including modified immunoglobulins and cleavage products
CC that can be useful in gene therapy and the generation of transgenic
CC plants and animals. Furthermore, the CDR mimetibody is useful for
CC preparing compositions for modulating, treating or reducing the symptoms
CC of immune, cardiovascular, infectious, malignant and/or neurologic
CC diseases, as well as anaemia. Accordingly, they exhibit immunomodulator,
CC cardiant, antimicrobial, cytostatic and neuroprotective activities. This
CC peptide sequence is an SH3 antagonist peptide sequence used to make a
CC mimetibody of the invention.
XX
SQ Sequence 12 AA;
Query Match 97.9%; Score 47; DB 7; Length 12;
Best Local Similarity 100.0%; Pred. No. 8.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 RPLPPLXP 12
Db |||||
4 RPLPPLXP 12
RESULT 14
ID ADJ73398 standard; peptide; 12 AA.
XX
AC ADJ73398;
XX
DT 06-MAY-2004 (first entry)
DE SH3 antagonist peptide sequence SeqID 853.
XX
XX mimetic; CDR mimetibody; gene therapy; transgenic; immune;
KW cardiovascular; infectious; malignant; neurologic disease; anaemia;
KW immunomodulator; cardiant; antimicrobial; cytostatic; neuroprotective;
KW SH3.
XX Synthetic.
OS
XX WO2003084477-A2.
XX
XX 16-OCT-2003.
XX
XX 24-MAR-2003; 2003WO-US009139.
XX
XX 29-MAR-2002; 2002US-0368791P.
XX
XX (CENZ) CENTOCOR INC.
XX
XX Heavner GA, Knight DM, Scallion BJ, Ghayeb J;
XX WPI; 2003-804237/75.
XX
XX New CDR mimetibody comprising a portion of a heavy or light chain
PT variable region comprising human framework or ligand binding region,
PT useful for preparing a composition for treating e.g., immune,
PT cardiovascular or neurologic disease.
XX Disclosure; SEQ ID NO 853; 97pp; English.
PS

XX This invention relates to novel mammalian CDR mimetibodies, specific
CC portions or variants thereof. Specifically, it refers to an antibody
CC fragment where a protein has been inserted into, or replaces a portion
CC of, one or more CDR regions, such that each CDR mimetibody comprises at
CC least one portion of a heavy chain or light chain variable region, which
CC itself comprises at least one human framework region and at least one
CC ligand binding region (LBR). The present invention describes human
CC mimetibodies, including modified immunoglobulins and cleavage products
CC that can be useful in gene therapy and the generation of transgenic
CC plants and animals. Furthermore, the CDR mimetibody is useful for
CC preparing compositions for modulating, treating or reducing the symptoms
CC of immune, cardiovascular, infectious, malignant and/or neurologic
CC diseases, as well as anaemia. Accordingly, they exhibit immunomodulator,
CC cardiant, antimicrobial, cytostatic and neuroprotective activities. This
CC peptide sequence is an SH3 antagonist peptide sequence used to make a
CC mimetibody of the invention.
XX
SQ Sequence 12 AA;
Query Match 97.9%; Score 47; DB 7; Length 12;
Best Local Similarity 100.0%; Pred. No. 8.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 RPLPPLXP 12
Db |||||
4 RPLPPLXP 12
RESULT 15
ID ADJ53032 standard; peptide; 12 AA.
XX
AC ADJ53032;
XX
DT 06-MAY-2004 (first entry)
DE CH1 deleted mimetibody-related peptide SeqID853.
XX
XX CH1 deleted mimetibody; immunosuppressive; cardiovascular; cardiant;
KW hypotensive; neuroprotective; nootropic; antibacterial; virucide;
KW fungicide; gene therapy; immune disorder; cardiovascular disease;
KW arrhythmia; hypertension; heart failure; neurodegenerative;
KW multiple sclerosis; dementia; Alzheimer's disease; anaemia;
KW cancerous condition; infectious disease; bacterial infection;
KW viral infection; fungal infection.
XX
XX Unidentified.
OS Synthetic.
XX
XX Key Location/Qualifiers
FT Misc-difference 1..12 /note= "All Xaa's in this sequence are unidentified amino
FT acids"
XX
XX WO2004002417-A2.
XX
XX 08-JAN-2004.
XX
XX 27-JUN-2003; 2003WO-US020347.
XX
XX 28-JUN-2002; 2002US-0392431P.
XX
XX (CENZ) CENTOCOR INC.
XX
XX Heavner GA, Knight DM, Ghayeb J, Scallion BJ, Nesspor TC;
XX Kutoloski KA;
XX WPI; 2004-082870/08.
XX
XX New CH1-deleted mimetibody polypeptides and nucleic acids, useful for
PT modulating, treating, alleviating, preventing an immune, cardiovascular,
PT or neurodegenerative disease or disorder, anemia, cancer, or infectious
PT

```

PT diseases.
XX
PS Claim 3; SEQ ID NO 853; 129pp; English.
XX
CC This invention relates to CHI deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an immunosuppressive,
CC cardiovascular, cardiant, hypotensive, neuroprotective, nootropic,
CC antibacterial, virucide or fungicide activity. In addition, the disclosed
CC sequences may prove useful for gene therapy. The CHI-deleted mimetibody
CC is useful for diagnosing or treating a disease condition in a cell,
CC tissue, organ or animal, specifically for modulating, treating,
CC alleviating, preventing the incidence or reducing the symptoms of an
CC immune, cardiovascular (for example arrhythmia, hypertension or heart
CC failure), or neurodegenerative (for example multiple sclerosis, dementia
CC or Alzheimer's disease) diseases or disorders, anaemia, cancerous
CC conditions, or infectious diseases (for example bacterial, viral or
CC fungal infection). The present sequence is that of a peptide which may be
XX used during the creation of a mimetibody of the invention.
XX
SQ Sequence 12 AA;
Query Match 97.9%; Score 47; DB 8; Length 12;
Best Local Similarity 100.0%; Pred. No. 8.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 4 RPLPPLXP 12
Db 4 RPLPPLXP 12
|||||
|||||

RESULT 16
ADJ53035
ID ADJ53035 standard; peptide; 12 AA.
XX
AC ADJ53035;
XX
DT 06-MAY-2004 (first entry)
XX
DE CHI deleted mimetibody-related peptide SeqID856.
XX
KW CHI deleted mimetibody; immunosuppressive; cardiovascular; cardiant;
KW hypotensive; neuroprotective; nootropic; antibacterial; virucide;
KW fungicide; gene therapy; immune disorder; cardiovascular disease;
KW arrhythmia; hypertension; heart failure; neurodegenerative;
KW multiple sclerosis; dementia; Alzheimer's disease; anaemia;
KW cancerous condition; infectious disease; bacterial infection;
KW viral infection; fungal infection.
XX
OS Unidentified.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 1..12
FT /note= "All Xaa's in this sequence are unidentified amino
FT acids"
XX
XX WO2004002417-A2.
XX
XX 08-JAN-2004.
XX
XX 27-JUN-2003; 2003WO-US020347.
XX
XX 28-JUN-2002; 2002US-0392431P.
XX
XX (CENZ ) CENTOCOR INC.
XX
XX Heavner GA, Knight DM, Ghrayeb J, Scallion BJ, Nesspor TC;
XX Kutloski KA;
XX WPI; 2004-082870/08.
XX
XX New CHI-deleted mimetibody polypeptides and nucleic acids, useful for

```

```

PT modulating, treating, alleviating, preventing an immune, cardiovascular,
PT or neurodegenerative disease or disorder, anemia, cancer, or infectious
PT diseases.
XX
PS Claim 3; SEQ ID NO 856; 129pp; English.
XX
CC This invention relates to CHI deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an immunosuppressive,
CC cardiovascular, cardiant, hypotensive, neuroprotective, nootropic,
CC antibacterial, virucide or fungicide activity. In addition, the disclosed
CC sequences may prove useful for gene therapy. The CHI-deleted mimetibody
CC is useful for diagnosing or treating a disease condition in a cell,
CC tissue, organ or animal, specifically for modulating, treating,
CC alleviating, preventing the incidence or reducing the symptoms of an
CC immune, cardiovascular (for example arrhythmia, hypertension or heart
CC failure), or neurodegenerative (for example multiple sclerosis, dementia
CC or Alzheimer's disease) diseases or disorders, anaemia, cancerous
CC conditions, or infectious diseases (for example bacterial, viral or
CC fungal infection). The present sequence is that of a peptide which may be
XX used during the creation of a mimetibody of the invention.
XX
SQ Sequence 12 AA;
Query Match 97.9%; Score 47; DB 8; Length 12;
Best Local Similarity 100.0%; Pred. No. 8.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 4 RPLPPLXP 12
Db 4 RPLPPLXP 12
|||||
|||||

RESULT 17
ADJ51993
ID ADJ51993 standard; peptide; 12 AA.
XX
AC ADJ51993;
XX
DT 06-MAY-2004 (first entry)
XX
DE CHI deleted mimetibody-related peptide SeqID853.
XX
KW CHI deleted mimetibody; osteopathic; cardiovascular-Gen;
KW dermatological-Gen; auditory; endocrine-Gen; gastrointestinal-Gen;
KW gynaecological-Gen; hepatotropic; haemostatic; immunomodulator;
KW antiallergic; muscular-Gen; cytostatic; antiinflammatory; neuroleptic;
KW ophthalmological; nephrotropic; respiratory-Gen; tumour necrosis factor;
KW TNF; cytokine; bone disorder; joint disorder; cardiovascular disorder;
KW dental disorder; oral disorder; dermatological disorder; ear disorder;
KW nose disorder; throat disorder; endocrine disorder; metabolic disorder;
KW gastrointestinal disorder; gynaecological disorder; hepatic disorder;
KW obstetric disorder; haematological disorder; immunological disorder;
KW allergic disorder; infectious disorder; musculoskeletal disorder;
KW oncological disorder; neurological disorder; nutritional disorder;
KW ophthalmologic disorder; pediatric disorder; psychiatric disorder;
KW renal disorder; pulmonary disorder.
XX
OS Unidentified.
OS Synthetic.
XX
XX
XX Key Location/Qualifiers
FT Misc-difference 1..12
FT /note= "All Xaa's in this sequence are unidentified amino
FT acids"
XX
XX WO2004002424-A2.
XX
XX 08-JAN-2004.
XX
XX 30-JUN-2003; 2003WO-US020495.
XX
XX 28-JUN-2002; 2002US-0392431P.

```

```

PR 19-SEP-2002; 2002US-0412144P.
XX (CENZ ) CENTOCOR INC.
XX
XX Heavner GA, Knight DM, Ghrayeb J, Scallion BJ, Neaspor TC;
XX Kutoloski KA;
XX WPI; 2004-082872/08.
XX
XX New CHI deleted mimetibody polypeptide and nucleic acid, useful for
XX diagnosing, preventing or treating cardiovascular, dermatologic,
XX endocrine, gastrointestinal, gynecologic, infectious, neurologic and
XX nutritional disorders.
XX
XX Claim 15; SEQ ID NO 853; 123pp; English.
XX
XX This invention relates to CHI deleted mimetibodies (and the DNA sequences
XX which encode them), compositions, methods and uses. The invention may be
XX useful for the development of compounds with an osteopathic,
XX cardiovascular-Gen, dermatological-Gen, auditory, endocrine-Gen,
XX gastrointestinal-Gen, gynaecological-Gen, hepatotropic, haemostatic,
XX immunomodulator, anti-allergic, muscular-Gen, cytostatic,
XX anti-inflammatory, neuroleptic, ophthalmological, nephrotropic or
XX respiratory-Gen activity acting as a tumour necrosis factor (TNF)-
XX modulator or cytokine-agonist. The methods and compositions of the
XX present invention are useful for the diagnosis, prevention and/or
XX treatment of diseases or conditions associated with aberrant expression
XX or activity of the CHI deleted mimetibody, such as a bone or joint,
XX cardiovascular, dental or oral, dermatological, ear, nose or throat,
XX endocrine, metabolic, gastrointestinal, gynaecological, hepatic,
XX obstetric, haematologic, immunological, allergic, infectious,
XX musculoskeletal, oncological, neurological, nutritional, ophthalmologic,
XX pediatric, psychiatric, renal or pulmonary disorders. The present
XX sequence is that of a peptide which may be used during the creation of a
XX mimetibody of the invention.
XX
XX Sequence 12 AA;
XX
XX Query Match 97.9%; Score 47; DB 8; Length 12;
XX Best Local Similarity 100.0%; Pred. No. 8.5;
XX Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 4 RPLPPLPXP 12
XX Db |||||
XX 4 RPLPPLPXP 12
XX
XX RESULT 18
XX ADJ51996
XX ID ADJ51996 standard; peptide; 12 AA.
XX
XX AC ADJ51996;
XX
XX XX
XX DT 06-MAY-2004 (first entry)
XX
XX DE CHI deleted mimetibody-related peptide seqID856.
XX
XX CHI deleted mimetibody; osteopathic; cardiovascular-Gen;
XX dermatological-Gen; auditory; endocrine-Gen; gastrointestinal-Gen;
XX gynaecological-Gen; hepatotropic; haemostatic; immunomodulator;
XX anti-allergic; muscular-Gen; cytostatic; anti-inflammatory; neuroleptic;
XX ophthalmological; nephrotropic; respiratory-Gen; tumour necrosis factor;
XX TNF; cytokine; bone disorder; joint disorder; cardiovascular disorder;
XX dental disorder; oral disorder; dermatological disorder; ear disorder;
XX nose disorder; throat disorder; endocrine disorder; metabolic disorder;
XX gastrointestinal disorder; gynaecological disorder; hepatic disorder;
XX obstetric disorder; haematologic disorder; immunologic disorder;
XX allergic disorder; infectious disorder; musculoskeletal disorder;
XX oncological disorder; neurological disorder; nutritional disorder;
XX ophthalmologic disorder; pediatric disorder; psychiatric disorder;
XX renal disorder; pulmonary disorder.
XX
XX Unidentified.
XX
OS

```

OS Synthetic.

XX Key Location/Qualifiers

XX Misc-difference 1..12 /note= "All Xaa's in this sequence are unidentified amino acids"

XX WO2004002424-A2.

XX 08-JAN-2004.

XX 30-JUN-2003; 2003WO-US020495.

XX 28-JUN-2002; 2002US-0392431P.

XX 19-SEP-2002; 2002US-0412144P.

XX (CENZ) CENTOCOR INC.

XX Heavner GA, Knight DM, Ghrayeb J, Scallion BJ, Neaspor TC; Kutoloski KA;

XX WPI; 2004-082872/08.

XX New CHI deleted mimetibody polypeptide and nucleic acid, useful for diagnosing, preventing or treating cardiovascular, dermatologic, endocrine, gastrointestinal, gynecologic, infectious, neurologic and nutritional disorders.

XX Claim 15; SEQ ID NO 856; 123pp; English.

XX This invention relates to CHI deleted mimetibodies (and the DNA sequences which encode them), compositions, methods and uses. The invention may be useful for the development of compounds with an osteopathic, cardiovascular-Gen, dermatological-Gen, auditory, endocrine-Gen, gastrointestinal-Gen, gynaecological-Gen, hepatotropic, haemostatic, immunomodulator, anti-allergic, muscular-Gen, cytostatic, anti-inflammatory, neuroleptic, ophthalmological, nephrotropic or respiratory-Gen activity acting as a tumour necrosis factor (TNF)-modulator or cytokine-agonist. The methods and compositions of the present invention are useful for the diagnosis, prevention and/or treatment of diseases or conditions associated with aberrant expression or activity of the CHI deleted mimetibody, such as a bone or joint, cardiovascular, dental or oral, dermatological, ear, nose or throat, endocrine, metabolic, gastrointestinal, gynaecological, hepatic, obstetric, haematologic, immunological, allergic, infectious, musculoskeletal, oncological, neurological, nutritional, ophthalmologic, pediatric, psychiatric, renal or pulmonary disorders. The present sequence is that of a peptide which may be used during the creation of a mimetibody of the invention.

XX Sequence 12 AA;

XX Query Match 97.9%; Score 47; DB 8; Length 12;

XX Best Local Similarity 100.0%; Pred. No. 8.5;

XX Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX QY 4 RPLPPLPXP 12

XX Db |||||

XX 4 RPLPPLPXP 12

XX RESULT 19

XX AAR93461

XX ID AAR93461 standard; peptide; 15 AA.

XX AC AAR93461;

XX XX 09-MAY-1996 (first entry)

XX DE GST-PI3K protein tyrosine kinase derived peptide #1.

XX SH3 ligand; SH3 binding agent; biased phage library;

XX recognition sequence; src SH3 domain; Paget's disease; restenosis;

KW rheumatoid arthritis; gout; neutrophil oxidase; tyrosine kinase; p47;
 KW p67; complex; chronic myelogenous leukaemia; cancer.
 XX Synthetic.
 XX WO9524419-A1.
 XX PD 14-SEP-1995.
 XX PF 13-MAR-1995; 95WO-US003208.
 XX PR 11-MAR-1994; 94US-00209835.
 XX PR 06-JAN-1995; 95US-00369832.
 XX PA (ARIA-) ARIAD PHARM INC.
 XX PI Rickles RJ, Brugge JS, Botfield MC, Zoller MJ;
 XX WPI; 1995-328231/42.
 XX PT Identification of peptide(s) binding specifically to SH3 domains - for
 PT use in inhibiting interactions mediated by SH3 domains in treatment of
 PT e.g. osteoporosis and cancer.
 XX Disclosure; Fig 5; 74pp; English.
 CC The sequences given in AAR93457-71 represent peptides which are SH3
 CC ligands/SH3 binding agents. They represent a biased phage library which
 CC comprises five random amino acids flanking the decapeptide -RSRPLPLP or
 CC derivatives of this, which was identified as a recognition sequence for
 CC the src SH3 domain. These sequences were identified using the method of
 CC the invention. The method comprises contacting the SH3 domain with a
 CC mixture of peptides under conditions permitting a ligand to bind to an
 CC SH3 domain to form a complex. Any unbound peptides are removed and the
 CC complexed peptide ligands are dissociated from the complexes. The
 CC selected peptides are enriched by re-contacting them with the SH3 domain
 CC and then candidates which bind to the SH3 domain are detected. The
 CC isolated SH3 binding peptides may be used in the diagnosis, prevention
 CC and treatment of conditions or diseases resulting from cellular processes
 CC mediated by an SH3-based interaction. Such diseases include Paget's
 CC disease. Other conditions treatable with these peptides include
 CC reutenosis, rheumatoid arthritis, gout and other problems in which an SH3
 CC of neutrophil oxidase p47 and p67 complex is implicated, etc
 XX Sequence 15 AA;
 SQ Query Match 97.9%; Score 47; DB 2; Length 15;
 Best Local Similarity 88.9%; Pred. No. 10;
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 4 RPLPPLPXP 12
 DB |||||
 4 RPLPPLPFP 12
 RESULT 20
 AAR93457
 ID AAR93457 standard; peptide; 15 AA.
 XX AC AAR93457;
 XX DT 09-MAY-1996 (first entry)
 XX GST-LYN protein tyrosine kinase derived peptide #1.
 XX SH3 ligand; SH3 binding agent; biased phage library;
 KW recognition sequence; src SH3 domain; Paget's disease; reutenosis;
 KW rheumatoid arthritis; gout; neutrophil oxidase; tyrosine kinase; p47;
 KW p67; complex; chronic myelogenous leukaemia; cancer.
 XX Synthetic.
 XX WO9524419-A1.

XX 14-SEP-1995.
 XX PF 13-MAR-1995; 95WO-US003208.
 XX PR 11-MAR-1994; 94US-00209835.
 XX PR 06-JAN-1995; 95US-00369832.
 XX PA (ARIA-) ARIAD PHARM INC.
 XX PI Rickles RJ, Brugge JS, Botfield MC, Zoller MJ;
 XX WPI; 1995-328231/42.
 XX PT Identification of peptide(s) binding specifically to SH3 domains - for
 PT use in inhibiting interactions mediated by SH3 domains in treatment of
 PT e.g. osteoporosis and cancer.
 XX Disclosure; Fig 5; 74pp; English.
 CC The sequences given in AAR93457-71 represent peptides which are SH3
 CC ligands/SH3 binding agents. They represent a biased phage library which
 CC comprises five random amino acids flanking the decapeptide -RSRPLPLP or
 CC derivatives of this, which was identified as a recognition sequence for
 CC the src SH3 domain. These sequences were identified using the method of
 CC the invention. The method comprises contacting the SH3 domain with a
 CC mixture of peptides under conditions permitting a ligand to bind to an
 CC SH3 domain to form a complex. Any unbound peptides are removed and the
 CC complexed peptide ligands are dissociated from the complexes. The
 CC selected peptides are enriched by re-contacting them with the SH3 domain
 CC and then candidates which bind to the SH3 domain are detected. The
 CC isolated SH3 binding peptides may be used in the diagnosis, prevention
 CC and treatment of conditions or diseases resulting from cellular processes
 CC mediated by an SH3-based interaction. Such diseases include Paget's
 CC disease. Other conditions treatable with these peptides include
 CC reutenosis, rheumatoid arthritis, gout and other problems in which an SH3
 CC of neutrophil oxidase p47 and p67 complex is implicated, etc
 XX Sequence 15 AA;
 SQ Query Match 97.9%; Score 47; DB 2; Length 15;
 Best Local Similarity 88.9%; Pred. No. 10;
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 4 RPLPPLPXP 12
 DB |||||
 4 RPLPPLPFP 12
 RESULT 21
 AAR93460
 ID AAR93460 standard; peptide; 15 AA.
 XX AC AAR93460;
 XX DT 09-MAY-1996 (first entry)
 XX GST-LYN protein tyrosine kinase derived peptide #4.
 XX SH3 ligand; SH3 binding agent; biased phage library;
 KW recognition sequence; src SH3 domain; Paget's disease; reutenosis;
 KW rheumatoid arthritis; gout; neutrophil oxidase; tyrosine kinase; p47;
 KW p67; complex; chronic myelogenous leukaemia; cancer.
 XX Synthetic.
 XX WO9524419-A1.
 XX PD 14-SEP-1995.
 XX PF 13-MAR-1995; 95WO-US003208.
 XX PR 11-MAR-1994; 94US-00209835.

```

PR 06-JAN-1995; 95US-00369832.
XX (ARIA-) ARIAD PHARM INC.
XX Rickles RJ, Brugge JS, Botfield MC, Zoller MJ;
XX WPI; 1995-328231/42.
DR
XX
XX Identification of peptide(s) binding specifically to SH3 domains - for
PT use in inhibiting interactions mediated by SH3 domains in treatment of
PT e.g. osteoporosis and cancer.
XX
XX Disclosure; Fig 5; 74pp; English.
XX
XX The sequences given in AAR93457-71 represent peptides which are SH3
CC ligands/SH3 binding agents. They represent a biased phage library which
CC comprises five random amino acids flanking the decapeptide -RSRLPPLP or
CC derivatives of this, which was identified as a recognition sequence for
CC the src SH3 domain. These sequences were identified using the method of
CC the invention. The method comprises contacting the SH3 domain with a
CC mixture of peptides under conditions permitting a ligand to bind to an
CC SH3 domain to form a complex. Any unbound peptides are removed and the
CC complexed peptide ligands are dissociated from the complexes. The
CC selected peptides are enriched by re-contacting them with the SH3 domain
CC and then candidates which bind to the SH3 domain are detected. The
CC isolated SH3 binding peptides may be used in the diagnosis, prevention
CC and treatment of conditions or diseases resulting from cellular processes
CC mediated by an SH3-based interaction. Such diseases include Paget's
CC disease. Other conditions treatable with these peptides include
CC restenosis, rheumatoid arthritis, gout and other problems in which an SH3
CC of neutrophil oxidase p47 and p67 complex is implicated, etc
XX
XX Sequence 15 AA;
SQ
Query Match 97.9%; Score 47; DB 2; Length 15;
Best Local Similarity 88.9%; Pred. No. 10;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 RPLPPLXP 12
DB ||||| |
4 RPLPPLPLP 12

RESULT 22
AAR93458
ID AAR93458 standard; peptide; 15 AA.
XX
XX AAR93458;
XX
XX 09-MAY-1996 (first entry)
XX
XX GST-LYN protein tyrosine kinase derived peptide #2.
XX
XX SH3 ligand; SH3 binding agent; biased phage library;
XX recognition sequence; src SH3 domain; Paget's disease; restenosis;
XX rheumatoid arthritis; gout; neutrophil oxidase; tyrosine kinase; p47;
XX p67; complex; chronic myelogenous leukaemia; cancer.
XX
XX Synthetic.
XX
XX WO9524419-A1.
XX
XX 14-SEP-1995.
XX
XX 13-MAR-1995; 95WO-US003208.
XX
XX 11-MAR-1994; 94US-00209835.
XX
XX 06-JAN-1995; 95US-00369832.
XX
XX (ARIA-) ARIAD PHARM INC.
XX
XX Rickles RJ, Brugge JS, Botfield MC, Zoller MJ;
XX WPI; 1995-328231/42.
XX
XX Identification of peptide(s) binding specifically to SH3 domains - for
PT use in inhibiting interactions mediated by SH3 domains in treatment of
PT e.g. osteoporosis and cancer.
XX
XX

```


PS Disclosure; Fig 5; 74pp; English.

XX The sequences given in AAR93457-71 represent peptides which are SH3

CC ligands/SH3 binding agents. They represent a biased phage library which

CC comprises five random amino acids flanking the decapeptide -RSRPLPLP or

CC derivatives of this, which was identified as a recognition sequence for

CC the src SH3 domain. These sequences were identified using the method of

CC the invention. The method comprises contacting the SH3 domain with a

CC mixture of peptides under conditions permitting a ligand to bind to an

CC SH3 domain to form a complex. Any unbound peptides are removed and the

CC complexed peptide ligands are dissociated from the complexes. The

CC selected peptides are enriched by re-contacting them with the SH3 domain

CC and then candidates which bind to the SH3 domain are detected. The

CC isolated SH3 binding peptides may be used in the diagnosis, prevention

CC and treatment of conditions or diseases resulting from cellular processes

CC mediated by an SH3-based interaction. Such diseases include Paget's

CC disease. Other conditions treatable with these peptides include

CC restenosis, rheumatoid arthritis, gout and other problems in which an SH3

CC of neutrophil oxidase p47 and p67 complex is implicated, etc

XX

SQ Sequence 15 AA;

Query Match 97.9%; Score 47; DB 2; Length 15;

Best Local Similarity 88.9%; Pred. No. 10;

Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 RPLPPLPXP 12

Db 4 RPLPPLP 12

RESULT 24

AAW05451

ID AAW05451 standard; peptide; 15 AA.

XX

AC AAW05451;

XX

DT 24-FEB-1998 (first entry)

XX

DE SH3-binding peptide bSH3002.

XX

KW Src-homology region 3 domain; human; mouse; SH3 domain; cell growth;

KW cellular signalling element; cellular structural element; malignancy;

KW protein identification; functional domain; protein screening;

KW cellular signal transduction process; binding peptide.

XX

OS Synthetic.

XX

PN WO9631625-A1.

XX

PD 10-OCT-1996.

XX

PF 04-APR-1996; 96WO-US004454.

XX

PR 07-APR-1995; 95US-00417872.

PR 03-APR-1996; 96US-00630915.

XX

PA (CYTO-) CYTOGEN CORP.

PA (UYNC-) UNIV NORTH CAROLINA.

XX

PI Sparks AB, Hoffman N, Kay BK, Fowlkes DM, McConnell SJ;

XX

DR WPI; 1996-465045/46.

XX

PT Identifying polypeptide(s) having specific functional domain (esp. SH3

PT domain) - comprises detecting selective binding to recognition unit,

PT regardless of sequence homology.

XX

PS Example; Fig 12A; 174pp; English.

XX

CC AAW05445-W05492 represent Src-homology region 3 (SH2) domain binding

CC peptides. These sequences were used as parts of multivalent recognition

CC unit complexes used in the method of the invention. The method of the

CC

invention is for identifying polypeptides containing functional domains

of interest (especially SH3 domains). It comprises contacting a

multivalent recognition unit (RU) complex with a number of peptides and

identifying polypeptides having a selective binding affinity for the RU

complex. The method is based on functional similarities and does not rely

on sequence similarities. Prior methods only gave limited success for

identifying proteins containing an SH3 domain due to the minimal sequence

homology among known SH3 proteins. Multivalent RU complexes are

particularly suited to screening for polypeptides containing functional

domains that are similar to, but not identical in sequence to, the

original target functional domain. The new method enables proteins having

a common function to be identified. Identification of novel SH3 proteins

will be useful for a better understanding of cell growth, malignancy,

signal transduction processes, etc. New candidate drugs can be

identified, and their specificities (e.g. pharmacological activities) can

be assessed using the method of the invention

XX

SQ Sequence 15 AA;

Query Match 97.9%; Score 47; DB 2; Length 15;

Best Local Similarity 88.9%; Pred. No. 10;

Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 RPLPPLPXP 12

Db 5 RPLPPLPDP 13

RESULT 25

AAW37659

ID AAW37659 standard; peptide; 15 AA.

XX

AC AAW37659;

XX

DT 23-APR-1998 (first entry)

XX

DE PPPY motif containing peptide bSH3002 used to bind WW domains.

XX

KW Peptide recognition unit; WW domain; cell signalling; growth regulation;

KW cytoskeleton organisation; targeted drug screening; modulator;

KW WW domain interaction; YAP protein; dystrophin.

XX

OS Synthetic.

XX

PN WO9737223-A1.

XX

PD 09-OCT-1997.

XX

PF 03-APR-1997; 97WO-US005547.

XX

PR 03-APR-1996; 96US-00630916.

XX

PA (CYTO-) CYTOGEN CORP.

PA (UYNC-) UNIV NORTH CAROLINA.

XX

PI Pirozzi G, Kay BK, Fowlkes DM;

XX

DR WPI; 1997-503234/46.

XX

PT Identifying cell signalling and growth regulatory polypeptides by

PT reaction with multivalent recognition complex - polypeptides are useful

PT in targeted drug selection.

XX

PS Example 6.3; Fig 7; 220pp; English.

XX

CC Peptides AAW37653-77 contain PPPY-like motifs. The PPPY motif is found

CC in the proline rich regions of WBP-1 and WBP-2 proteins. Peptides

CC containing this residue have been shown to bind the YAP WW domain, but

CC not the WW domain from dystrophin or to a panel of SH3 domains. Peptides

CC AAW37653-77 were biotinylated and complexed with alkaline streptavidin,

CC and used in a cross affinity mapping experiment. They were tested for

CC their ability to bind to the 12 individual novel WW domains of WWP1

CC (AAW36794), WWP2 (AAW36795), WWP3 (AAW37696) and WWP4 (AAW36797), which

CC were expressed as glutathione-S-transferase expression proteins. The present peptide, derived from a potassium channel, does not bind to WW domains of the novel proteins. The WW domain is a small functional domain. Its name is derived from the observation that two tryptophan residues, one in the amino terminal portion of the WW domain and one in the carboxyl terminal portion, are conserved. Most proteins containing WW domains have a function involving cell signalling and growth regulation or the organisation of the cytoskeleton. Polypeptides containing a WW domain are identified by treating a multivalent recognition unit complex that has selective binding affinity for a WW domain, with many polypeptides and identifying those with selective affinity for the complex. Proteins containing WW domains are used for targeted drug screening, i.e. to identify potential modulators of specific WW domain interactions

XX
SQ Sequence 15 AA;

Query Match 97.9%; Score 47; DB 2; Length 15;
Best Local Similarity 88.9%; Pred. No. 10;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 RPLPPLPXP 12
Db 5 RPLPPLPDP 13
|||||

RESULT 26
AAW38924
ID AAW38924 standard; peptide; 15 AA.
XX
AC AAW38924;
XX
XX
DT 27-MAR-1998 (first entry)
XX
DE Peptide resembling an SH3 domain binding peptide SEQ ID NO:321.
XX
XX Cortactin; SH3 domain; binding peptide; Src homology region 3;
KW tyrosine kinase; immune response; lymphokine; interleukin 1; Nck; Abl;
KW PLCgamma; p53bp2; Crk; Yes; Grb2.
XX
XX Synthetic.
OS
XX WO9730074-A1.
PN
XX 21-AUG-1997.
PD
XX 14-FEB-1997; 97WO-US002298.
PF
XX 16-FEB-1996; 96US-00602999.
PR
XX (CYTO-) CYTOGEN CORP.
XX (UYNC-) UNIV NORTH CAROLINA.
PA
XX Sparks AB, Kay BK, Thorn JM, Quilliam LA, Der CJ, Fowlkes DM;
PI Rider JE;
XX WPI; 1997-424972/39.
DR
XX Src homology region 3 binding peptide - used to activate Src tyrosine kinase(s) and to stimulate immune response by increasing production of certain lymphokine(s), e.g. interleukin-1.
PT
PT
PT
PS
PS Claim 22; Page 90; 131pp; English.
XX
XX The present sequence represents a peptide which resembles a Src homology region 3 (SH3) binding peptide. SH3 binding peptides are selected from: (a) peptides which bind the SH3 domain of Cortactin; (b) peptides which bind the middle SH3 domain of Nck; (c) peptides which bind the SH3 domain of Abl; (d) peptides which bind the SH3 domain of Src; (e) peptides which bind the SH3 domain of PLC gamma; (f) peptides which bind the SH3 domain of p53bp2; (g) peptides which bind the amino-terminal SH3 domain of Crk; (h) peptides which bind the SH3 domain of Yes; and (i) peptides which bind the amino-terminal SH3 domain of Grb2. The purified binding peptides

CC can be used in the method to identify inhibitors of their binding to their respective SH3 domains, which could be used to modulate the pharmacological activity of proteins or polypeptide containing the SH3 domain. The peptides can also be used to activate Src or Src-related protein tyrosine kinases, to stimulate the immune response by increasing the production of certain lymphokines, e.g. tumour necrosis factor-alpha and interleukin-1, or to deliver a conjugated molecule to certain cellular compartments containing Src or Src related proteins

XX
SQ Sequence 15 AA;

Query Match 97.9%; Score 47; DB 2; Length 15;
Best Local Similarity 88.9%; Pred. No. 10;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 RPLPPLPXP 12
Db 5 RPLPPLPDP 13
|||||

RESULT 27
ADB49285
ID ADB49285 standard; peptide; 15 AA.
XX
AC ADB49285;
XX
DT 04-DEC-2003 (first entry)
XX
DE Novel WW domain binding peptide #7.
XX
XX WW domain; drug candidate screening; drug discovery; drug modification;
KW drug refinement; immunogen; WW binding protein; WW domain.
XX
XX Unidentified.
OS
XX US2003077577-A1.
PN
XX 24-APR-2003.
PD
XX 28-JUN-2002; 2002US-00185050.
PF
XX 03-APR-1996; 96US-00630916.
PR
XX 03-APR-1997; 97US-00826516.
PR
XX (PIRO/) PIROZZI G.
PA (KAYB/) KAY B K.
PA (FOWL/) FOWLKES D M.
XX
XX Pirozzi G, Kay BK, Fowlkes DM;
PI WPI; 2003-635075/60.
DR
XX Novel purified polypeptide comprising WW domain, useful for drug discovery, modification and refinement, for discovering polypeptides involved in pharmacological activities, or as an immunogen to generate antibodies.
PT
PT
PT
PS
PS Example; Fig 7; 133pp; English.
XX
XX The invention describes a purified polypeptide (I) comprising a WW domain which has a sequence (S1) selected from 11 sequences fully defined in the specification, a sequence (S2) selected from 48 sequences fully defined in the specification or a sequence (S3) comprising 683, 906, 924 or 725 amino acids fully defined in the specification. (I) is useful for screening a potential drug candidate, by allowing (I) to come into contact with at least one recognition unit having a selective affinity for the WW domain in (I), in the presence of an amount of a potential drug candidate, such that (I) and the recognition unit are capable of interacting when brought into contact with one another in the absence of the drug candidate, and determining the effect, if any, of the presence of the amount of the drug candidate on the interaction of (I) with the recognition unit. (I) is useful for drug discovery, modification and refinement, for discovering polypeptides involved in pharmacological

CC activities, or as an immunogen to generate antibodies. This is the amino
CC acid sequence of a peptide that binds the novel WW domains of the
CC invention

XX Sequence 15 AA;
SQ Query Match 97.9%; Score 47; DB 7; Length 15;
Best Local Similarity 88.9%; Pred. No. 10;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 RPLPPLXP 12
Db ||||| |
5 RPLPPLPDP 13

RESULT 28
AAB17217
ID AAB17217 standard; peptide; 17 AA.
XX
AC AAB17217;
XX
DT 31-OCT-2000 (first entry)
XX
DE Mast cell antagonist/Mast cell protease inhibitor peptide SEQ ID NO:273.
XX
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
KW thrombosis; pharmaceutical.
XX
OS Synthetic.
XX
PN WO200024782-A2.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US025044.
XX
PR 23-OCT-1998; 98US-0105371P.
PR 22-OCT-1999; 99US-00428082.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham J, Boone TC;
XX WPI; 2000-350702/30.
XX
PT Novel composition of matter comprising an Fc domain and pharmacologically
PT active peptides, useful for treating cancer and autoimmune diseases.
XX
PS Claim 39; Page 292; 608pp; English.
XX
CC The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
CC (L2)d-P2-(L3)e-P3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
CC P3, and P4 = are each independently sequences of pharmacologically active
CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,
CC c, d, e, and f = are each independently 0 or 1, provided that at least 1
CC of a and b is 1. The composition can have cytostatic, antiasthmatic,
CC thrombolytic and immunosuppressive activities. DNAs, vectors and host
CC cells from the present invention can be used for producing pharmaceutical
CC compositions. The compositions are useful for treating cancer, asthma,
CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
CC a Fab domain) can provide a longer half-life or incorporate functions
CC such as Fc receptor binding, protein A binding, complement fixation, and
CC possibly placental transfer. AAG69443 to AAG69526 and AAB16955 to
CC AAB18003 represent nucleotide and amino acid sequences used in the
CC exemplification of the present invention

XX Sequence 17 AA;
SQ Query Match 97.9%; Score 47; DB 3; Length 17;
Best Local Similarity 88.9%; Pred. No. 12;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 RPLPPLXP 12
Db ||||| |
6 RPLPPLPLP 14

RESULT 29
ABB73208
ID ABB73208 standard; peptide; 17 AA.
XX
AC ABB73208;
XX
DT 05-APR-2002 (first entry)
XX
DE Mast cell antagonist/mast cell protease inhibitor peptide SEQ ID NO:273.
XX
KW Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TWP;
KW TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;
KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
KW cytostatic; anorectic; antiarthritic; antiinfertility; antidiabetic; ophthalmological;
KW antianemic; anorectic; antiinfertility; haemostatic; dermatological;
KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
KW sleep disorder; neurological degenerative disease; anaemia;
KW thrombocytopaenia; metastatic tumour; systemic lupus erythematosus;
KW Fanconi's syndrome.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200183525-A2.
XX
PD 08-NOV-2001.
XX
PF 02-MAY-2001; 2001WO-US014310.
XX
PR 03-MAY-2000; 2000US-00563286.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;
XX WPI; 2002-130313/17.
XX
PT Novel vehicle-peptide molecule or its multimers useful for treating
PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
PT diabetic retinopathy, obesity, sleep disorders and infertility.
XX
PS Claim 39; Page 54; 176pp; English.
XX
CC The present invention describes a vehicle-peptide molecule (I) or its
CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,
CC cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,
CC antianemic, anorectic, antiinfertility, haemostatic, dermatological and
CC neuroprotective activities. (I) can be used as a therapeutic or
CC prophylactic agent as well as for screening purposes. (I) is useful for
CC diagnosing diseases characterised by dysfunction of their associated
CC protein of interest, for identifying normal or abnormal proteins of
CC interest, as a part of diagnostic kit to detect the presence of their
CC proteins of interest in a biological sample. Additionally, (I) is useful
CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,
CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,
CC infertility, and neurological degenerative diseases. (I), comprising EPO-
CC mimetic compounds are useful for treating disorders characterised by low
CC red blood cell levels such as anaemia. The TPO-mimetic comprising


```

Query Match      97.9%; Score 47; DB 8; Length 17;
Best Local Similarity 88.9%; Pred. No. 12;
Matches      8; Conservative      0; Mismatches      1; Indels      0; Gaps      0;

QY      4 RPLPPLXP 12
      |||||
DB      6 RPLPPLPLP 14

RESULT 32
ADJ51957
ID      ADJ51957 standard; peptide; 17 AA.
AC
XX
AC      ADJ51957;
XX
DT      06-MAY-2004 (first entry)
XX
XX      CH1 deleted mimetibody-related peptide SeqID816.
XX
XX      CH1 deleted mimetibody; osteopathic; cardiovascular-Gen;
KW      dermatological-Gen; auditory; endocrine-Gen; gastrointestinal-Gen;
KW      gynaecological-Gen; hepatotropic; haemostatic; immunomodulator;
KW      anti-allergic; muscular-Gen; cytostatic; anti-inflammatory; neuroleptic;
KW      ophthalmological; nephrotropic; respiratory-Gen; tumour necrosis factor;
KW      TNF; cytokine; bone disorder; joint disorder; cardiovascular disorder;
KW      dental disorder; oral disorder; dermatological disorder; ear disorder;
KW      nose disorder; throat disorder; endocrine disorder; metabolic disorder;
KW      gastrointestinal disorder; gynaecological disorder; hepatic disorder;
KW      obstetric disorder; haematological disorder; immunological disorder;
KW      allergic disorder; infectious disorder; musculoskeletal disorder;
KW      oncological disorder; neurological disorder; nutritional disorder;
KW      ophthalmological disorder; pediatric disorder; psychiatric disorder;
KW      renal disorder; pulmonary disorder.
XX
XX      Unidentified.
OS
XX      Synthetic.
XX
XX      W02004002424-A2.
PN
XX
XX      08-JAN-2004.
PD
XX
XX      30-JUN-2003; 2003WO-US020495.
PF
XX
XX      28-JUN-2002; 2002US-0392431P.
PR
XX      19-SEP-2002; 2002US-0412144P.
XX
XX      (CENZ ) CENTOCOR INC.
PA
XX
XX      Heavner GA, Knight DM, Chrayeb J, Scallan BJ, Nesspor TC;
PI      Kutoloski KA;
XX
XX      WPI; 2004-082872/08.
DR
XX
XX      New CHI deleted mimetibody polypeptide and nucleic acid, useful for
PT      diagnosing, preventing or treating cardiovascular, dermatologic,
PT      endocrine, gastrointestinal, gynecologic, infectious, neurologic and
PT      nutritional disorders.
XX
XX      Claim 15; SEQ ID NO 816; 123pp; English.
PS
XX
XX      This invention relates to CHI deleted mimetibodies (and the DNA sequences
CC      which encode them), compositions, methods and uses. The invention may be
CC      useful for the development of compounds with an osteopathic,
CC      cardiovascular-Gen, dermatological-Gen, auditory, endocrine-Gen,
CC      gastrointestinal-Gen, gynaecological-Gen, hepatotropic, haemostatic,
CC      immunomodulator, anti-allergic, muscular-Gen, cytostatic,
CC      anti-inflammatory, neuroleptic, ophthalmological, nephrotropic or
CC      respiratory-Gen activity acting as a tumour necrosis factor (TNF) -
CC      modulator or cytokine-agonist. The methods and compositions of the
CC      present invention are useful for the diagnosis, prevention and/or
CC      treatment of diseases or conditions associated with aberrant expression
CC      or activity of the CHI deleted mimetibody, such as a bone or joint,
CC      cardiovascular, dental or oral, dermatological, ear, nose or throat,

```

```

CC      endocrine, metabolic, gastrointestinal, gynaecological, hepatic,
CC      obstetric, haematologic, immunological, allergic, infectious,
CC      musculoskeletal, oncological, neurological, nutritional, ophthalmologic,
CC      pediatric, psychiatric, renal or pulmonary disorders. The present
CC      sequence is that of a peptide which may be used during the creation of a
CC      mimetibody of the invention.
XX
SQ      Sequence 17 AA;

Query Match      97.9%; Score 47; DB 8; Length 17;
Best Local Similarity 88.9%; Pred. No. 12;
Matches      8; Conservative      0; Mismatches      1; Indels      0; Gaps      0;

QY      4 RPLPPLXP 12
      |||||
DB      6 RPLPPLPLP 14

RESULT 33
AAY41645
ID      AAY41645 standard; peptide; 22 AA.
XX
AC      AAY41645;
XX
DT      02-DEC-1999 (first entry)
XX
DE      Human peptide hkv1.5-62-83.
XX
XX      SH3 domain; binding motif; potassium channel; protein tyrosine kinase;
KW      proline rich.
XX
XX      Homo sapiens.
OS
XX      US5955259-A.
XX
XX      21-SEP-1999.
PD
XX
XX      19-DEC-1996; 96US-00769745.
PF
XX
XX      19-DEC-1996; 96US-00769745.
PR
XX      (UYBR-) UNIV BRANDEIS.
XX
XX      Holmes TC, Levitan IB;
PI
XX
XX      WPI; 1999-560490/47.
DR
XX
XX      Identification of compounds that modulate potassium ion channel binding
PT      with protein tyrosine kinase SH3 domains.
XX
XX      Example 2; Col 14; 18pp; English.
XX
XX      A method has been developed for determining if a compound modulates the
CC      binding of a potassium ion channel to the SH3 domain of a protein
CC      tyrosine kinase by contacting the channel with a polypeptide comprising
CC      the SH3 domain and the compound to be assessed and measuring channel-SH3
CC      binding. The method is useful for assessing the ability of a compound to
CC      modulate the formation of channel-SH3 domain complexes to improve the
CC      understanding of mechanisms of potassium channel blockage and assess the
CC      ability of potential therapeutics to inhibit blockage. The present
CC      sequence represents the Human peptide hkv1.5-62-83, which is used in an
CC      example from the present invention
XX
XX      Sequence 22 AA;

Query Match      97.9%; Score 47; DB 2; Length 22;
Best Local Similarity 88.9%; Pred. No. 15;
Matches      8; Conservative      0; Mismatches      1; Indels      0; Gaps      0;

QY      4 RPLPPLXP 12
      |||||
DB      4 RPLPPLPLP 12

```

```

RESULT 34
AAAY41612
ID AAY41612 standard; peptide; 25 AA.
XX
AC AAY41612;
XX
DT 02-DEC-1999 (first entry)
XX
DE Mammalian ion channel proline rich motif containing peptide #6.
XX
KW SH3 domain; binding motif; potassium channel; protein tyrosine kinase;
XX proline rich.
XX
OS Homo sapiens.
XX
PN US9595259-A.
XX
PD 21-SEP-1999.
XX
PF 19-DEC-1996; 96US-00769745.
XX
PR 19-DEC-1996; 96US-00769745.
XX
PA (UYBR-) UNIV BRANDEIS.
XX
PI Holmes TC, Levitan IB;
XX
DR WPI; 1999-560490/47.
XX
PT Identification of compounds that modulate potassium ion channel binding
XX with protein tyrosine kinase SH3 domains.
XX
PS Disclosure; Col 8; 18pp; English.
XX
CC A method has been developed for determining if a compound modulates the
XX binding of a potassium ion channel to the SH3 domain of a protein
XX tyrosine kinase by contacting the channel with a polypeptide comprising
XX the SH3 domain and the compound to be assessed and measuring channel-SH3
XX binding. The method is useful for assessing the ability of a compound to
XX modulate the formation of channel-SH3 domain complexes to improve the
XX understanding of mechanisms of potassium channel blockage and assess the
XX ability of potential therapeutics to inhibit blockage. AAY41607 to
XX AAY41644 represent mammalian ion channel peptides with proline-rich
XX motifs
XX
SQ Sequence 25 AA;
Query Match 97.9%; Score 47; DB 2; Length 25;
Best Local Similarity 88.9%; Pred. No. 16;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 4 RPLPPLPXP 12
DB 1 RPLPPLPDP 9
|||||||
1 RPLPPLPDP 9

RESULT 35
ABR42419
ID ABR42419 standard; protein; 613 AA.
XX
AC ABR42419;
XX
DT 11-AUG-2003 (first entry)
XX
DE Human potassium channel Kv1.5.
XX
KW Human; potassium channel; Kv1.5; neuronal disease; multiple sclerosis;
XX neuroprotective.
XX
OS Homo sapiens.
XX
PN WO2003035690-A2.

```

```

XX
PD 01-MAY-2003.
XX
PF 25-OCT-2002; 2002WO-GB004821.
XX
PR 25-OCT-2001; 2001GB-00025636.
XX
PA (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.
XX
PI Dolly JO, Akhtar S, Shamotienko O;
XX
DR WPI; 2003-421404/39.
XX
PT Preparing a K+ channel fusion protein comprising alpha subunits and beta
XX subunits, useful in the manufacture of a medicament for treating multiple
XX sclerosis.
XX
PS Disclosure; Fig 4; 93pp; English.
XX
CC The present sequence is the protein sequence of human potassium channel
XX Kv1.5. The invention provides oligomeric potassium channel fusion
XX proteins, preferably comprising Kv1.1 and/or Kv1.2 subunits, and methods
XX of producing them in mammalian cells. The fusion proteins can be used to
XX identify compounds that modulate potassium channel activity. Such
XX modulator compounds are useful for the treatment of multiple sclerosis
XX
SQ Sequence 613 AA;
Query Match 97.9%; Score 47; DB 6; Length 613;
Best Local Similarity 88.9%; Pred. No. 2.9e+02;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 4 RPLPPLPXP 12
DB 65 RPLPPLPDP 73
|||||||

RESULT 36
AAR93362
ID AAR93362 standard; peptide; 12 AA.
XX
AC AAR93362;
XX
DT 24-APR-1996 (first entry)
XX
DE LYN/PI3K protein tyrosine kinase derived peptide.
XX
KW SH3 ligand; SH3 binding agent; biased phage library;
XX recognition sequence; src SH3 domain; Paget's disease; restenosis;
XX rheumatoid arthritis; gout; neutrophil oxidase; tyrosine kinase; p47;
XX p67; complex; chronic myelogenous leukaemia; cancer.
XX
OS Synthetic.
XX
PN WO9524419-A1.
XX
PD 14-SEP-1995.
XX
PF 13-MAR-1995; 95WO-US0003208.
XX
PR 11-MAR-1994; 94US-00209835.
XX
PR 06-JAN-1995; 95US-00369832.
XX
PA (ARIA-) ARIAD PHARM INC.
XX
PI Rickles RJ, Brugge JS, Botfield MC, Zoller MJ;
XX
DR WPI; 1995-328231/42.
XX
PT Identification of peptide(s) binding specifically to SH3 domains - for
XX use in inhibiting interactions mediated by SH3 domains in treatment of
XX e.g. osteoporosis and cancer.

```

PS Disclosure; Fig 2; 74pp; English.

The sequences given in AAR93343-68 represent peptides which are SH3 ligands/SH3 binding agents. They represent a biased phage library which comprises six random amino acids flanking the hexapeptide RSRPRP- which was identified as a recognition sequence for the src SH3 domain. These sequences were identified using the method of the invention. The method comprises contacting the SH3 domain with a mixture of peptides under conditions permitting a ligand to bind to an SH3 domain to form a complex. Any unbound peptides are removed and the complexed peptide ligands are dissociated from the complexes. The selected peptides are enriched by re-contacting them with the SH3 domain and then candidates which bind to the SH3 domain are detected. The isolated SH3 binding peptides may be used in the diagnosis, prevention and treatment of conditions or diseases resulting from cellular processes mediated by an SH3-based interaction. Such diseases include Paget's disease. Other conditions treatable with these peptides include retinosis, rheumatoid arthritis, gout and other problems in which an SH3 of neutrophil oxidase p47 and p67 complex is implicated etc

Sequence 12 AA;

Query Match	95.8%	Score 46;	DB 2;	Length 12;
Best Local Similarity	88.9%	Pred. No. 11;		
Matches	8;	Conservative	0;	Mismatches 1;
				Indels 0;
				Gaps 0;

Qy 4 RPLPPLXP 12
| | | | |
Db 4 RPLPPLPPP 12

RESULT 37

AAR93377
ID AAR93377 standard; peptide; 12 AA.

AA AAR93377:

DT 24-APR-1996 (first entry)

XX DE Grb-2 protein tyrosine kinase derived peptide #1.

SH3 ligand; SH3 binding agent; biased phage library;
recognition sequence; src SH3 domain; Paget's disease; restenosis;
rheumatoid arthritis; gout; neutrophil oxidase; tyrosine kinase; p47;
p67; complex; chronic myelogenous leukaemia; cancer.

OS Synthetic.

XX PN WO9524419-A1.

14-SEP-1995
PD
XXXX
PF 13-MAR-1995: 95WO-US003208.11-MAR-1994: 94US-00209835-XX
PB

PR 06-JAN-1995; 95US-00369832.

PA (ARIA-) ARIAD PHARM INC.

PI Rickles RJ, Brugge JS, Botfield MC, Zoller MJ;

WPI: 1995-328231/42.

Identification of peptide(s) binding specifically to SH3 domains - for use in inhibiting interactions mediated by SH3 domains in treatment of e.g. osteoporosis and cancer.

PS Disclosure: Fig 2: 74pp: English.

xx CC The sequences given in MAR93377-80 represent peptides which are SH3
CC ligands/SH3 binding agents. They represent a biased phage library which
CC comprises six random amino acids flanking the hexapeptide KLYPL-
CC was identified as a recognition sequence for the src SH3 domain. These

sequences were identified using the method of the invention. The method comprises contacting the SH3 domain with a mixture of peptides under conditions permitting a ligand to bind to an SH3 domain to form a complex. Any unbound peptides are removed and the complexed peptide ligands are dissociated from the complexes. The selected peptides are enriched by re-contacting them with the SH3 domain and then candidates which bind to the SH3 domain are detected. The isolated SH3 binding peptides may be used in the diagnosis, prevention and treatment of conditions or diseases resulting from cellular processes mediated by an SH3-based interaction. Such diseases include Paget's disease. Other conditions treatable with these peptides include retinosis, rheumatoid arthritis, gout and other problems in which an SH3 of neutrophil oxidase p47 and p67 complex is implicated etc

Sequence 12 AA;

Query Match	95.8%	Score 46;	DB 2;	Length 12;
Best Local Similarity	88.9%	Pred. No. 11;		
Matches 8;	Conservative	0;	Mismatches	1; Indels
				0; Gaps
				0;

Qy 4 RPLPPLXP 12
| | | | |
Db 4 RPLPPLPW 12

RESULT 38

AAR93345
ID AAR93345 standard; peptide; 12 AA.

AAR93345:

24-APR-1996 (first entry)

DE SRC protein tyrosine kinase derived peptide #3.

SH3 ligand; SH3 binding agent; biased phage library;
recognition sequence; src SH3 domain; Paget's disease; restenosis;
rheumatoid arthritis; gout; neutrophil oxidase; tyrosine kinase; p47;
p57; complex; chronic myelogenous leukaemia; cancer.

OS Synthetic.

XX PN WO9524419-A1.

14-SEP-1995.

13-MAR-1995: 95WO-US003208.

11-MAR-1994: 94US-00209835-XX
PR

PR 06-JAN-1995; 95US-00369832.

PA (ARIA-) ARIAD PHARM INC.

PI Rickles RJ, Brugge JS, Botfield MC, Zoller MJ;

DR WPI; 1995-328231/42.

PT Identification of peptide(s) binding specifically to SH3 domains - for
PT use in inhibiting interactions mediated by SH3 domains in treatment of
PT e.g. osteoporosis and cancer.

PS Disclosure; Fig 2; 74pp; English.

The sequences given in AAR93343-68 represent peptides which are SH3 ligands/SH3 binding agents. They represent a biased phage library which comprises six random amino acids flanking the hexapeptide RSLRPL- which was identified as a recognition sequence for the src SH3 domain. These sequences were identified using the method of the invention. The method comprises contacting the SH3 domain with a mixture of peptides under conditions permitting a ligand to bind to an SH3 domain to form a complex. Any unbound peptides are removed and the complexed peptide ligands are dissociated from the complexes. The selected peptides are enriched by re-contacting them with the SH3 domain and then candidates

CC which bind to the SH3 domain are detected. The isolated SH3 binding peptides may be used in the diagnosis, prevention and treatment of CC conditions or diseases resulting from cellular processes mediated by an SH3-based interaction. Such diseases include Paget's disease. Other CC conditions treatable with these peptides include restenosis, rheumatoid arthritis, gout and other problems in which an SH3 of neutrophil oxidase CC p47 and p67 complex is implicated, etc

XX SQ Sequence 12 AA;

Query Match 95.8%; Score 46; DB 2; Length 12;
Best Local Similarity 88.9%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 RPLPPLPXP 12
|||||||

Db 4 RPLPPLPPP 12
|||||||

RESULT 39

AA17255

ID AAB17255 standard; peptide; 12 AA.

XX AAB17255;

AC AAB17255;

DT 31-OCT-2000 (first entry)

XX SH3 antagonist peptide sequence SEQ ID NO:311.

DE Modified peptide; therapeutic agent; fusion; Fc domain; cancer; autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF; immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP; inhibitor; erythropoietin; thrombopoietin; interleukin 1; cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor; vascular endothelial growth factor; matrix metalloproteinase; asthma; thrombosis; pharmaceutical.

XX Synthetic.

OS WO200024782-A2.

PN 04-MAY-2000.

PD 25-OCT-1999; 99WO-US025044.

PF 23-OCT-1998; 98US-0105371P.

PR 23-OCT-1999; 99US-00428082.

XX (AMGE-) AMGEN INC.

XX Feige U, Liu C, Cheetham J, Boone TC;

PI WPI; 2000-350702/30.

DR Novel composition of matter comprising an Fc domain and pharmacologically active peptides, useful for treating cancer and autoimmune diseases.

XX Claim 39; Page 304; 608pp; English.

XX The present invention describes composition of matter (I) comprising an Fc domain, pharmacologically active peptides, and linkers. Where (I) is: (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-P3, and P4 = are each independently sequences of pharmacologically active peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b, c, d, e, and f = are each independently 0 or 1, provided that at least 1 of a and b is 1. The composition can have cytostatic, antiasthmatic, thrombolytic and immunosuppressive activities. DNAs, vectors and host cells from the present invention can be used for producing pharmaceutical compositions. The compositions are useful for treating cancer, asthma, thrombosis, or autoimmune diseases. The use of an Fc domain (rather than a Fab domain) can provide a longer half-life or incorporate functions

CC such as Fc receptor binding, protein A binding, complement fixation, and CC possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to CC AAB18003 represent nucleotide and amino acid sequences used in the CC exemplification of the present invention

XX SQ Sequence 12 AA;

Query Match 95.8%; Score 46; DB 3; Length 12;
Best Local Similarity 88.9%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 RPLPPLPXP 12
|||||||

Db 4 RPLPPLPPP 12
|||||||

RESULT 40

ABB73248

ID ABB73248 standard; peptide; 12 AA.

XX ABB73248;

AC ABB73248;

XX 05-APR-2002 (first entry)

DT Src homology3 (SH3) antagonist peptide SEQ ID NO:311.

XX Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO; erythropoietin; TPO; tumour necrosis factor alpha inhibitor; TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP; TPO mimetic peptide; EPO mimetic peptide; BMP; VEGF antagonist; MMP inhibitor; antiinflammatory; antitumour; immunosuppressive; cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological; antianemic; anorectic; antiinfertility; haemostatic; dermatological; neuroprotective; inflammatory disease; autoimmune disease; tumour growth; cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity; sleep disorder; neurological degenerative disease; anaemia; thrombocytopaenia; metastatic tumour; systemic lupus erythematosus; Panconi's syndrome.

XX Homo sapiens.

OS Synthetic.

XX WO200183525-A2.

PN 08-NOV-2001.

PD 02-MAY-2001; 2001WO-US014310.

PF 03-MAY-2000; 2000US-00563286.

XX (AMGE-) AMGEN INC.

XX Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;

PI WPI; 2002-130313/17.

DR Novel vehicle-peptide molecule or its multimers useful for treating inflammatory and autoimmune diseases, cancer, rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders and infertility.

XX Claim 39; Page 55; 176pp; English.

XX The present invention describes a vehicle-peptide molecule (I) or its multimers. (I) can have antinflammatory, antitumour, immunosuppressive, cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological, antianemic, anorectic, antiinfertility, haemostatic, dermatological and neuroprotective activities. (I) can be used as a therapeutic or prophylactic agent as well as for screening purposes. (I) is useful for diagnosing diseases characterised by dysfunction of their associated protein of interest, for identifying normal or abnormal proteins of interest, as a part of diagnostic kit to detect the presence of their proteins of interest in a biological sample. Additionally, (I) is useful for treating inflammatory and autoimmune diseases, tumour growth, cancer,

CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,
 CC infertility, and neurological degenerative diseases. (1), comprising EPO-
 CC mimetic compounds are useful for treating disorders characterised by low
 CC red blood cell levels such as anaemia. The TPO-mimetic comprising
 CC compounds are useful for treating conditions that involve an existing
 CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet
 CC deficiency, such as thrombocytopenia, aplastic anaemia, metastatic
 CC tumour which result in thrombocytopenia, systemic lupus erythematosus,
 CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABLJ5695 to ABLJ5777
 CC represent amino acid and nucleic acid sequences used in the
 CC exemplification of the present invention

XX
 SQ Sequence 12 AA;

Query Match 95.8%; Score 46; DB 5; Length 12;
 Best Local Similarity 88.9%; Pred. No. 11;
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 RPLPPLPPX 12
 |||||
 Db 4 RPLPPLPPP 12

RESULT 41
 ADJ73402
 ID ADJ73402 standard; peptide; 12 AA.

XX AC ADJ73402;

XX DT 06-MAY-2004 (first entry)

XX DE SH3 antagonist peptide sequence SeqID 857.

XX KW mimetic; CDR mimetibody; gene therapy; transgenic; immune;
 KW cardiovascular; infectious; malignant; neurologic disease; anaemia;
 KW immunomodulator; cardiant; antimicrobial; cytostatic; neuroprotective;
 KW SH3.

XX OS Synthetic.

XX PN WO2003084477-A2.

XX PD 16-OCT-2003.

XX PF 24-MAR-2003; 2003WO-US009139.

XX PR 29-MAR-2002; 2002US-0368791P.

XX PA (CENZ) CENTOCOR INC.

XX PI Heavner GA, Knight DM, Scallion BJ, Ghayeb J;

XX DR WPI; 2003-804237/75.

XX PT New CDR mimetibody comprising a portion of a heavy or light chain
 PT variable region comprising human framework or ligand binding region,
 PT useful for preparing a composition for treating e.g., immune,
 PT cardiovascular or neurologic disease.

XX PS Disclosure; SEQ ID NO 857; 97pp; English.

XX CC This invention relates to novel mammalian CDR mimetibodies, specific
 CC portions or variants thereof. Specifically, it refers to an antibody
 CC fragment where a protein has been inserted into, or replaces a portion
 CC of, one or more CDR regions, such that each CDR mimetibody comprises at
 CC least one portion of a heavy chain or light chain variable region, which
 CC itself comprises at least one human framework region and at least one
 CC ligand binding region (LBR). The present invention describes human
 CC mimetibodies, including modified immunoglobulins and cleavage products
 CC that can be useful in gene therapy and the generation of transgenic
 CC plants and animals. Furthermore, the CDR mimetibody is useful for
 CC preparing compositions for modulating, treating or reducing the symptoms
 CC of immune, cardiovascular, infectious, malignant and/ or neurologic

CC diseases, as well as anaemia. Accordingly, they exhibit immunomodulator,
 CC cardiant, antimicrobial, cytostatic and neuroprotective activities. This
 CC peptide sequence is an SH3 antagonist peptide sequence used to make a
 CC mimetibody of the invention.

SQ Sequence 12 AA;

Query Match 95.8%; Score 46; DB 7; Length 12;
 Best Local Similarity 88.9%; Pred. No. 11;
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 RPLPPLPPX 12
 |||||
 Db 4 RPLPPLPPP 12

RESULT 42
 ADJ53036

ID ADJ53036 standard; peptide; 12 AA.

XX AC ADJ53036;

XX DT 06-MAY-2004 (first entry)

XX DE CHI deleted mimetibody-related peptide SeqID857.

XX KW CHI deleted mimetibody; immunosuppressive; cardiovascular; cardiant;
 KW hypotensive; neuroprotective; nootropic; antibacterial; virucide;
 KW fungicide; gene therapy; immune disorder; cardiovascular disease;
 KW arrhythmia; hypertension; heart failure; neurodegenerative;
 KW multiple sclerosis; dementia; Alzheimer's disease; anaemia;
 KW cancerous condition; infectious disease; bacterial infection;
 KW viral infection; fungal infection.

XX OS Unidentified.

XX OS Synthetic.

XX PH Key Location/Qualifiers

XX FT Misc-difference 1..12

XX FT /note= "All Xaa's in this sequence are unidentified amino
 acids"

XX PN WO2004002417-A2.

XX PD 08-JAN-2004.

XX PF 27-JUN-2003; 2003WO-US020347.

XX PR 28-JUN-2002; 2002US-0392431P.

XX PA (CENZ) CENTOCOR INC.

XX PI Heavner GA, Knight DM, Ghayeb J, Scallion BJ, Neesspor TC;

XX PI Kutoloski KA;

XX DR WPI; 2004-082870/08.

XX PT New CHI-deleted mimetibody polypeptides and nucleic acids, useful for
 PT modulating, treating, alleviating, preventing an immune, cardiovascular,
 PT or neurodegenerative disease or disorder, anemia, cancer, or infectious
 PT diseases.

XX PS Claim 3; SEQ ID NO 857; 129pp; English.

XX CC This invention relates to CHI deleted mimetibodies (and the DNA sequences
 CC which encode them), compositions, methods and uses. The invention may be
 CC useful for the development of compounds with an immunosuppressive,
 CC cardiovascular, cardiant, hypotensive, neuroprotective, nootropic,
 CC antibacterial, virucide or fungicide activity. In addition, the disclosed
 CC sequences may prove useful for gene therapy. The CHI-deleted mimetibody
 CC is useful for diagnosing or treating a disease condition in a cell,
 CC tissue, organ or animal, specifically for modulating, treating,
 CC alleviating, preventing the incidence or reducing the symptoms of an

CC immune, cardiovascular (for example arrhythmia, hypertension or heart
 CC failure), or neurodegenerative (for example multiple sclerosis, dementia
 CC or Alzheimer's disease) diseases or disorders, anaemia, cancerous
 CC conditions, or infectious diseases (for example bacterial, viral or
 CC fungal infection). The present sequence is that of a peptide which may be
 CC used during the creation of a mimetibody of the invention.
 XX
 SQ Sequence 12 AA;
 Query Match 95.8%; Score 46; DB 8; Length 12;
 Best Local Similarity 88.9%; Pred. NO. 11;
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 4 RPLPPLPXP 12
 Db |||||
 4 RPLPPLPPP 12
 RESULT 43
 ADJ51997
 ID ADJ51997 standard; peptide; 12 AA.
 XX
 AC ADJ51997;
 DT 06-MAY-2004 (first entry)
 XX
 DE CHI deleted mimetibody-related peptide SeqID857.
 XX
 KW CHI deleted mimetibody; osteopathic; cardiovascular-Gen;
 KW dermatological-Gen; auditory; endocrine-Gen; gastrointestinal-Gen;
 KW gynaecological-Gen; hepatotropic; haemostatic; immunomodulator;
 KW anti-allergic; muscular-Gen; cytotoxic; anti-inflammatory; neuroleptic;
 KW ophthalmological; nephrotropic; respiratory-Gen; tumour necrosis factor;
 KW TNF; cytokine; bone disorder; joint disorder; cardiovascular disorder;
 KW dental disorder; oral disorder; dermatological disorder; ear disorder;
 KW nose disorder; throat disorder; endocrine disorder; metabolic disorder;
 KW gastrointestinal disorder; gynaecological disorder; hepatic disorder;
 KW obstetric disorder; haematologic disorder; immunological disorder;
 KW allergic disorder; infectious disorder; musculoskeletal disorder;
 KW oncological disorder; neurological disorder; nutritional disorder;
 KW ophthalmologic disorder; pediatric disorder; psychiatric disorder;
 KW renal disorder; pulmonary disorder.
 XX
 OS Unidentified.
 OS Synthetic.
 XX
 PH Key Location/Qualifiers
 FT Misc-difference 1..12
 FT /note="All Xaa's in this sequence are unidentified amino
 FT acids"
 XX
 PN W02004002424-A2.
 XX
 XX 08-JAN-2004.
 XX
 XX 30-JUN-2003; 2003WO-US020495.
 XX
 PR 28-JUN-2002; 2002US-0392431P.
 PR 19-SEP-2002; 2002US-0412144P.
 XX
 XX (CENZ) CENTOCOR INC.
 XX
 XX Heavner GA, Knight DM, Ghayeb J, Scallion BJ, Nesspor TC;
 PI Kutloski KA;
 XX
 XX WPI; 2004-082872/08.
 DR
 XX New CHI deleted mimetibody polypeptide and nucleic acid, useful for
 PT diagnosing, preventing or treating cardiovascular, dermatologic, and
 PT endocrine, gastrointestinal, gynecologic, infectious, neurologic and
 PT nutritional disorders.
 XX
 XX Claim 15; SEQ ID NO 857; 123pp; English.

XX This invention relates to CHI deleted mimetibodies (and the DNA sequences
 CC which encode them), compositions, methods and uses. The invention may be
 CC useful for the development of compounds with an osteopathic-Gen,
 CC cardiovascular-Gen, dermatological-Gen, auditory, endocrine-Gen,
 CC gastrointestinal-Gen, gynaecological-Gen, hepatotropic, haemostatic,
 CC immunomodulator, anti-allergic, muscular-Gen, cytotoxic,
 CC anti-inflammatory, neuroleptic, ophthalmological, nephrotropic or
 CC respiratory-Gen activity acting as a tumour necrosis factor (TNF)-
 CC present invention are useful for the diagnosis, prevention and/or
 CC treatment of diseases or conditions associated with aberrant expression
 CC or activity of the CHI deleted mimetibody, such as a bone or joint,
 CC cardiovascular, dental or oral, dermatological, ear, nose or throat,
 CC endocrine, metabolic, gastrointestinal, gynaecological, hepatic,
 CC obstetric, haematologic, immunological, allergic, infectious,
 CC musculoskeletal, oncological, neurological, nutritional, ophthalmologic,
 CC pediatric, psychiatric, renal or pulmonary disorders. The present
 CC sequence is that of a peptide which may be used during the creation of a
 CC mimetibody of the invention.
 XX
 SQ Sequence 12 AA;
 Query Match 95.8%; Score 46; DB 8; Length 12;
 Best Local Similarity 88.9%; Pred. NO. 11;
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 4 RPLPPLPXP 12
 Db |||||
 4 RPLPPLPPP 12
 RESULT 44
 AAR93386
 ID AAR93386 standard; peptide; 14 AA.
 XX
 AC AAR93386;
 XX
 DT 30-APR-1996 (first entry)
 XX
 DE GST-SRC SH3 protein tyrosine kinase derived peptide #6.
 XX
 KW SH3 ligand; SH3 binding agent; biased phage library;
 KW recognition sequence; src SH3 domain; Paget's disease; restenosis;
 KW rheumatoid arthritis; gout; neutrophil oxidase; tyrosine kinase; p47;
 KW p67; complex; chronic myelogenous leukaemia; cancer.
 XX
 OS Synthetic.
 XX
 PN W09524419-A1.
 XX
 PD 14-SEP-1995.
 XX
 XX 13-MAR-1995; 95WO-US003208.
 XX
 XX 11-MAR-1994; 94US-00209835.
 PR 06-JAN-1995; 95US-00369832.
 XX
 XX (ARIA-) ARIAD PHARM INC.
 XX
 XX Rickles RJ, Brugge JS, Botfield MC, Zoller MJ;
 XX WPI; 1995-328231/42.
 DR
 XX Identification of peptide(s) binding specifically to SH3 domains - for
 PT use in inhibiting interactions mediated by SH3 domains in treatment of
 PT e.g. osteoporosis and cancer.
 XX
 PS Disclosure; Fig 4; 74pp; English.
 XX
 CC The sequences given in AAR93381-443 represent peptides which are SH3
 CC ligands/SH3 binding agents. They represent a biased phage library which
 CC comprises five random amino acids flanking the nonapeptide -RPLPPLPPP

CC which was identified as a recognition sequence for the src SH3 domain.
 CC These sequences were identified using the method of the invention. The
 CC method comprises contacting the SH3 domain with a mixture of peptides
 CC under conditions permitting a ligand to bind to an SH3 domain to form a
 CC complex. Any unbound peptides are removed and the complexed peptide
 CC ligands are dissociated from the complexes. The selected peptides are
 CC enriched by re-contacting them with the SH3 domain and then candidates
 CC which bind to the SH3 domain are detected. The isolated SH3 binding
 CC peptides may be used in the diagnosis, prevention and treatment of
 CC conditions or diseases resulting from cellular processes mediated by an
 CC SH3-based interaction. Such diseases include Paget's disease. Other
 CC conditions treatable with these peptides include restenosis, rheumatoid
 CC arthritis, gout and other problems in which an SH3 of neutrophil oxidase
 CC p47 and p67 complex is implicated, etc

XX SQ Sequence 14 AA;

Query Match 95.8%; Score 46; DB 2; Length 14;
 Best Local Similarity 88.9%; Pred. No. 13;
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 4 RPLPPLPXP 12
 |||||
 Db 6 RPLPPLPPP 14

RESULT 45

AAR93404
 ID AAR93404 standard; peptide; 14 AA.

XX AC AAR93404;

XX DT 30-APR-1996 (first entry)

XX GST-FYN SH3 protein tyrosine kinase derived peptide #14.

XX SH3 ligand; SH3 binding agent; biased phage library;
 KW recognition sequence; src SH3 domain; Paget's disease; restenosis;
 KW rheumatoid arthritis; gout; neutrophil oxidase; tyrosine kinase; p47;
 KW p67; complex; chronic myelogenous leukaemia; cancer.

XX OS Synthetic.

XX PN WO9524419-A1.

XX PD 14-SEP-1995.

XX PF 13-MAR-1995; 95WO-US003208.

XX PR 11-MAR-1994; 94US-00209835.

XX PR 06-JAN-1995; 95US-00369832.

XX PA (ARIA-) ARIAD PHARM INC.

XX PI Rickles RJ, Brugge JS, Botfield MC, Zoller MJ;

XX WPI; 1995-328231/42.

XX Identification of peptide(s) binding specifically to SH3 domains - for
 PT use in inhibiting interactions mediated by SH3 domains in treatment of
 PT e.g. osteoporosis and cancer.

XX Disclosure; Fig 4; 74pp; English.

XX The sequences given in AAR93381-443 represent peptides which are SH3
 CC ligands/SH3 binding agents. They represent a biased phage library which
 CC comprises five random amino acids flanking the nonapeptide -RPLPPLPPP
 CC which was identified as a recognition sequence for the src SH3 domain.
 CC These sequences were identified using the method of the invention. The
 CC method comprises contacting the SH3 domain with a mixture of peptides
 CC under conditions permitting a ligand to bind to an SH3 domain to form a
 CC complex. Any unbound peptides are removed and the complexed peptide
 CC ligands are dissociated from the complexes. The selected peptides are

CC enriched by re-contacting them with the SH3 domain and then candidates
 CC which bind to the SH3 domain are detected. The isolated SH3 binding
 CC peptides may be used in the diagnosis, prevention and treatment of
 CC conditions or diseases resulting from cellular processes mediated by an
 CC SH3-based interaction. Such diseases include Paget's disease. Other
 CC conditions treatable with these peptides include restenosis, rheumatoid
 CC arthritis, gout and other problems in which an SH3 of neutrophil oxidase
 CC p47 and p67 complex is implicated, etc

XX SQ Sequence 14 AA;

Query Match 95.8%; Score 46; DB 2; Length 14;
 Best Local Similarity 88.9%; Pred. No. 13;
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 4 RPLPPLPXP 12
 |||||
 Db 6 RPLPPLPPP 14

RESULT 46

AAR93450
 ID AAR93450 standard; peptide; 14 AA.

XX AC AAR93450;

XX DT 30-APR-1996 (first entry)

XX GST-PI3K SH3 protein tyrosine kinase derived peptide #7.

XX SH3 ligand; SH3 binding agent; biased phage library;
 KW recognition sequence; src SH3 domain; Paget's disease; restenosis;
 KW rheumatoid arthritis; gout; neutrophil oxidase; tyrosine kinase; p47;
 KW p67; complex; chronic myelogenous leukaemia; cancer.

XX OS Synthetic.

XX PN WO9524419-A1.

XX PD 14-SEP-1995.

XX PF 13-MAR-1995; 95WO-US003208.

XX PR 11-MAR-1994; 94US-00209835.

XX PR 06-JAN-1995; 95US-00369832.

XX PA (ARIA-) ARIAD PHARM INC.

XX PI Rickles RJ, Brugge JS, Botfield MC, Zoller MJ;

XX WPI; 1995-328231/42.

XX Identification of peptide(s) binding specifically to SH3 domains - for
 PT use in inhibiting interactions mediated by SH3 domains in treatment of
 PT e.g. osteoporosis and cancer.

XX Disclosure; Fig 4; 74pp; English.

XX The sequences given in AAR93444-56 represent peptides which are SH3
 CC ligands/SH3 binding agents. They represent a biased phage library which
 CC comprises five random amino acids flanking the nonapeptide -RPLPPLPPP
 CC which was identified as a recognition sequence for the src SH3 domain.
 CC These sequences were identified using the method of the invention. The
 CC method comprises contacting the SH3 domain with a mixture of peptides
 CC under conditions permitting a ligand to bind to an SH3 domain to form a
 CC complex. Any unbound peptides are removed and the complexed peptide
 CC ligands are dissociated from the complexes. The selected peptides are
 CC enriched by re-contacting them with the SH3 domain and then candidates
 CC which bind to the SH3 domain are detected. The isolated SH3 binding
 CC peptides may be used in the diagnosis, prevention and treatment of
 CC conditions or diseases resulting from cellular processes mediated by an
 CC SH3-based interaction. Such diseases include Paget's disease. Other
 CC conditions treatable with these peptides include restenosis, rheumatoid

CC arthritis, gout and other problems in which an SH3 of neutrophil oxidase
 CC p47 and p67 complex is implicated, etc
 XX
 SQ Sequence 14 AA;

Query Match 95.8%; Score 46; DB 2; Length 14;
 Best Local Similarity 88.9%; Pred. No. 13;
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 RPLPLPLXP 12
 DB 6 RPLPLPLPPP 14
 |||||

RESULT 47
 AAR93388
 ID AAR93388 standard; peptide; 14 AA.
 XX
 AC AAR93388;
 XX
 DT 30-APR-1996 (first entry)
 XX
 DE GST-SRC SH3 protein tyrosine kinase derived peptide #8.
 XX
 KW SH3 ligand; SH3 binding agent; biased phage library;
 KW recognition sequence; src SH3 domain; Paget's disease; restenosis;
 KW rheumatoid arthritis; gout; neutrophil oxidase; tyrosine kinase; p47;
 KW p67; complex; chronic myelogenous leukaemia; cancer.
 XX
 OS Synthetic.
 XX
 PN WO9524419-A1.
 XX
 PD 14-SEP-1995.
 XX
 PF 13-MAR-1995; 95WO-US003208.
 XX
 PR 11-MAR-1994; 94US-00209835.
 PR 06-JAN-1995; 95US-00369832.
 XX
 PA (ARIA-) ARIAD PHARM INC.
 XX
 PI Rickles RJ, Brugge JS, Botfield MC, Zoller MJ;
 XX WPI; 1995-328231/42.
 XX
 PT Identification of peptide(s) binding specifically to SH3 domains - for
 PT use in inhibiting interactions mediated by SH3 domains in treatment of
 PT e.g. osteoporosis and cancer.
 XX
 PS Disclosure; Fig 4; 74pp; English.
 XX
 CC The sequences given in AAR93381-443 represent peptides which are SH3
 CC ligands/SH3 binding agents. They represent a biased phage library which
 CC comprises five random amino acids flanking the nonapeptide -RPLPLPPP
 CC which was identified as a recognition sequence for the src SH3 domain.
 CC These sequences were identified using the method of the invention. The
 CC method comprises contacting the SH3 domain with a mixture of peptides
 CC under conditions permitting a ligand to bind to an SH3 domain to form a
 CC complex. Any unbound peptides are removed and the complexed peptide
 CC ligands are dissociated from the complexes. The selected peptides are
 CC enriched by re-contacting them with the SH3 domain and then candidates
 CC which bind to the SH3 domain are detected. The isolated SH3 binding
 CC peptides may be used in the diagnosis, prevention and treatment of
 CC conditions or diseases resulting from cellular processes mediated by an
 CC SH3-based interaction. Such diseases include Paget's disease. Other
 CC conditions treatable with these peptides include restenosis, rheumatoid
 CC arthritis, gout and other problems in which an SH3 of neutrophil oxidase
 CC p47 and p67 complex is implicated, etc
 XX
 SQ Sequence 14 AA;

Query Match 95.8%; Score 46; DB 2; Length 14;
 Best Local Similarity 88.9%; Pred. No. 13;
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 RPLPLPLXP 12
 DB 6 RPLPLPLPPP 14
 |||||

Best Local Similarity 88.9%; Pred. No. 13;
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 RPLPLPLXP 12
 DB 6 RPLPLPLPPP 14
 |||||

RESULT 48
 AAR93392
 ID AAR93392 standard; peptide; 14 AA.
 XX
 AC AAR93392;
 XX
 DT 30-APR-1996 (first entry)
 XX
 DE GST-PYN SH3 protein tyrosine kinase derived peptide #2.
 XX
 KW SH3 ligand; SH3 binding agent; biased phage library;
 KW recognition sequence; src SH3 domain; Paget's disease; restenosis;
 KW rheumatoid arthritis; gout; neutrophil oxidase; tyrosine kinase; p47;
 KW p67; complex; chronic myelogenous leukaemia; cancer.
 XX
 OS Synthetic.
 XX
 PN WO9524419-A1.
 XX
 PD 14-SEP-1995.
 XX
 PF 13-MAR-1995; 95WO-US003208.
 XX
 PR 11-MAR-1994; 94US-00209835.
 PR 06-JAN-1995; 95US-00369832.
 XX
 PA (ARIA-) ARIAD PHARM INC.
 XX
 PI Rickles RJ, Brugge JS, Botfield MC, Zoller MJ;
 XX WPI; 1995-328231/42.
 XX
 PT Identification of peptide(s) binding specifically to SH3 domains - for
 PT use in inhibiting interactions mediated by SH3 domains in treatment of
 PT e.g. osteoporosis and cancer.
 XX
 PS Disclosure; Fig 4; 74pp; English.
 XX
 CC The sequences given in AAR93381-443 represent peptides which are SH3
 CC ligands/SH3 binding agents. They represent a biased phage library which
 CC comprises five random amino acids flanking the nonapeptide -RPLPLPPP
 CC which was identified as a recognition sequence for the src SH3 domain.
 CC These sequences were identified using the method of the invention. The
 CC method comprises contacting the SH3 domain with a mixture of peptides
 CC under conditions permitting a ligand to bind to an SH3 domain to form a
 CC complex. Any unbound peptides are removed and the complexed peptide
 CC ligands are dissociated from the complexes. The selected peptides are
 CC enriched by re-contacting them with the SH3 domain and then candidates
 CC which bind to the SH3 domain are detected. The isolated SH3 binding
 CC peptides may be used in the diagnosis, prevention and treatment of
 CC conditions or diseases resulting from cellular processes mediated by an
 CC SH3-based interaction. Such diseases include Paget's disease. Other
 CC conditions treatable with these peptides include restenosis, rheumatoid
 CC arthritis, gout and other problems in which an SH3 of neutrophil oxidase
 CC p47 and p67 complex is implicated, etc
 XX
 SQ Sequence 14 AA;

Query Match 95.8%; Score 46; DB 2; Length 14;
 Best Local Similarity 88.9%; Pred. No. 13;
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 RPLPLPLXP 12
 DB 6 RPLPLPLPPP 14
 |||||

RESULT 49

AAR93400
ID AAR93400 standard; peptide; 14 AA.

XX AC AAR93400;
XX AC
XX DT 30-APR-1996 (first entry)
XX DE GST-FYN SH3 protein tyrosine kinase derived peptide #10.
XX KW SH3 ligand; SH3 binding agent; biased phage library;
XX KW recognition sequence; src SH3 domain; Paget's disease; restenosis;
XX KW rheumatoid arthritis; gout; neutrophil oxidase; tyrosine kinase; p47;
XX KW p67; complex; chronic myelogenous leukaemia; cancer.
XX OS Synthetic.
XX OS
XX PN WO9524419-A1.
XX PD 14-SEP-1995.
XX PF 13-MAR-1995; 95WO-US003208.
XX PR 11-MAR-1994; 94US-00209835.
XX PR 06-JAN-1995; 95US-00369832.
XX PA (ARIA-) ARIAD PHARM INC.
XX PI Rickles RJ, Brugge JS, Botfield MC, Zoller MJ;
XX WPI; 1995-328231/42.
XX

PT Identification of peptide(s) binding specifically to SH3 domains - for
PT use in inhibiting interactions mediated by SH3 domains in treatment of
PT e.g. osteoporosis and cancer.

PS Disclosure; Fig 4; 74pp; English.

XX The sequences given in AAR93381-443 represent peptides which are SH3
CC ligands/SH3 binding agents. They represent a biased phage library which
CC comprises five random amino acids flanking the nonapeptide -RPLPPLPPP
CC which was identified as a recognition sequence for the src SH3 domain.
CC These sequences were identified using the method of the invention. The
CC method comprises contacting the SH3 domain with a mixture of peptides
CC under conditions permitting a ligand to bind to an SH3 domain to form a
CC complex. Any unbound peptides are removed and the complexed peptide
CC ligands are dissociated from the complexes. The selected peptides are
CC enriched by re-contacting them with the SH3 domain and then candidates
CC which bind to the SH3 domain are detected. The isolated SH3 binding
CC peptides may be used in the diagnosis, prevention and treatment of
CC conditions or diseases resulting from cellular processes mediated by an
CC SH3-based interaction. Such diseases include Paget's disease. Other
CC conditions treatable with these peptides include restenosis, rheumatoid
CC arthritis, gout and other problems in which an SH3 of neutrophil oxidase
CC p47 and p67 complex is implicated, etc

SQ Sequence 14 AA;

Query Match 95.8%; Score 46; DB 2; Length 14;
Best Local Similarity 88.9%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 RPLPPLPXP 12
| | | | | | |
Db 6 RPLPPLPPP 14

RESULT 50

AAR93428
ID AAR93428 standard; peptide; 14 AA.

XX

AC AAR93428;

XX DT 30-APR-1996 (first entry)
XX DE GST-YES SH3 protein tyrosine kinase derived peptide #4.
XX KW SH3 ligand; SH3 binding agent; biased phage library;
XX KW recognition sequence; src SH3 domain; Paget's disease; restenosis;
XX KW rheumatoid arthritis; gout; neutrophil oxidase; tyrosine kinase; p47;
XX KW p67; complex; chronic myelogenous leukaemia; cancer.

XX OS Synthetic.

XX OS

XX PN WO9524419-A1.

XX PD 14-SEP-1995.

XX PF 13-MAR-1995; 95WO-US003208.

XX PR 11-MAR-1994; 94US-00209835.

XX PR 06-JAN-1995; 95US-00369832.

XX PA (ARIA-) ARIAD PHARM INC.

XX PI Rickles RJ, Brugge JS, Botfield MC, Zoller MJ;

XX WPI; 1995-328231/42.

XX Identification of peptide(s) binding specifically to SH3 domains - for
PT use in inhibiting interactions mediated by SH3 domains in treatment of
PT e.g. osteoporosis and cancer.
XX
XX Disclosure; Fig 4; 74pp; English.

XX The sequences given in AAR93381-443 represent peptides which are SH3
CC ligands/SH3 binding agents. They represent a biased phage library which
CC comprises five random amino acids flanking the nonapeptide -RPLPPLPPP
CC which was identified as a recognition sequence for the src SH3 domain.
CC These sequences were identified using the method of the invention. The
CC method comprises contacting the SH3 domain with a mixture of peptides
CC under conditions permitting a ligand to bind to an SH3 domain to form a
CC complex. Any unbound peptides are removed and the complexed peptide
CC ligands are dissociated from the complexes. The selected peptides are
CC enriched by re-contacting them with the SH3 domain and then candidates
CC which bind to the SH3 domain are detected. The isolated SH3 binding
CC peptides may be used in the diagnosis, prevention and treatment of
CC conditions or diseases resulting from cellular processes mediated by an
CC SH3-based interaction. Such diseases include Paget's disease. Other
CC conditions treatable with these peptides include restenosis, rheumatoid
CC arthritis, gout and other problems in which an SH3 of neutrophil oxidase
CC p47 and p67 complex is implicated, etc

SQ Sequence 14 AA;

Query Match 95.8%; Score 46; DB 2; Length 14;
Best Local Similarity 88.9%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 RPLPPLPXP 12
| | | | | | |
Db 6 RPLPPLPPP 14

Search completed: April 6, 2006, 09:30:42
Job time : 168.211 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 6, 2006, 09:35:13 ; Search time 21.7895 Seconds
(without alignments)
52.989 Million cell updates/sec

Title: US-10-632-388-307

Perfect score: 48

Sequence: 1 XXXRPLPLXP 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Database :

1: Pir1.*

2: Pir2.*

3: Pir3.*

4: Pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	47	97.9	613	2 A56031	potassium channel
2	43	89.6	162	2 T07173	hypothetical prote
3	42	87.5	198	2 E89008	protein W08A12.3 {
4	42	87.5	281	2 I38707	Fas ligand - human
5	42	87.5	473	2 S50755	hypothetical prote
6	41	85.4	339	2 JC7509	glycoprotein Vi-1
7	41	85.4	348	2 D88088	protein B0454.1 [i
8	41	85.4	459	2 S03116	gene 33 protein, h
9	41	85.4	477	2 T46917	hypothetical prote
10	41	85.4	535	2 A46101	protein-tyrosine-p
11	41	85.4	548	2 B46101	protein-tyrosine-p
12	41	85.4	627	2 T26064	hypothetical prote
13	41	85.4	663	1 TWMVR	protein-tyrosine k
14	41	85.4	894	2 F84870	hypothetical prote
15	41	85.4	1225	2 T16346	hypothetical prote
16	40	83.3	134	2 T36365	proline-rich prote
17	40	83.3	156	1 GNVQJ2	genome-linked prot
18	40	83.3	156	1 GNVQJL	genome-linked prot
19	40	83.3	156	1 GNVQWA	genome-linked prot
20	40	83.3	215	2 T15811	hypothetical prote
21	40	83.3	224	2 T16669	hypothetical prote
22	40	83.3	245	2 T29543	hypothetical prote
23	40	83.3	273	2 C70551	hypothetical prote
24	40	83.3	574	2 T41395	probable dna polym
25	40	83.3	775	2 C81594	hypothetical prote
26	40	83.3	775	2 D86549	hypothetical prote
27	40	83.3	775	2 B72074	hypothetical prote
28	40	83.3	793	1 S60735	splicing factor SF
29	39	81.2	93	2 G30010	hypothetical ORF-8

30	39	81.2	153	2 F83532	conserved hypothet
31	39	81.2	173	2 T19341	hypothetical prote
32	39	81.2	373	2 A70856	probable lppz prot
33	39	81.2	409	2 S60975	hypothetical prote
34	39	81.2	446	2 A34794	B-cell CLL/lymphom
35	39	81.2	472	2 A75464	probable zinc meta
36	39	81.2	514	2 C49507	potassium channel
37	39	81.2	602	2 A49507	potassium channel
38	39	81.2	832	2 T49494	condensin complex
39	39	81.2	1106	2 T31742	hypothetical prote
40	39	81.2	1211	2 T42230	AF4 protein - mou
41	39	81.2	1217	2 T42625	AF-4 protein - mou
42	39	81.2	1733	1 B45344	probable nuclear a
43	39	81.2	1958	2 B40505	hypothetical prote
44	39	81.2	2493	2 S72349	nonstructural poly
45	39	81.2	2493	2 S26372	nonstructural poly
46	38	79.2	191	2 E87601	OmpA family protei
47	38	79.2	217	2 T04353	DNA binding protei
48	38	79.2	222	2 C75539	conserved hypothet
49	38	79.2	308	2 A38582	pollen allergen pi
50	38	79.2	318	1 A31978	ADP,ATP carrier pr
51	38	79.2	336	2 E47301	Vir10 homolog - Bo
52	38	79.2	407	2 H87341	OmpA family protei
53	38	79.2	408	1 G65132	hypothetical 44.6
54	38	79.2	443	2 T49140	p62 ras-GAP associ
55	38	79.2	527	2 T84483	protein-tyrosine k
56	38	79.2	527	2 T49133	protein-tyrosine k
57	38	79.2	668	2 S56909	polymyxin B resist
58	38	79.2	925	2 S50490	hypothetical prote
59	38	79.2	1006	2 T42762	probable alpha-man
60	38	79.2	1386	2 T49316	profilaggrin relat
61	38	79.2	1560	2 T00080	hypothetical prote
62	37	77.1	68	2 F00050	synapcojanin 2 eps
63	37	77.1	79	2 T44145	B2 protein [import
64	37	77.1	101	2 T46506	hypothetical prote
65	37	77.1	104	2 D72542	hypothetical prote
66	37	77.1	109	2 G86433	protein T17H7.7 [i
67	37	77.1	124	2 G95982	hypothetical expor
68	37	77.1	161	2 T16207	hypothetical prote
69	37	77.1	166	2 T36616	hypothetical prote
70	37	77.1	175	2 S65828	probable movement
71	37	77.1	298	2 A75460	conserved hypothet
72	37	77.1	408	2 T32767	hypothetical prote
73	37	77.1	454	2 A75444	hypothetical prote
74	37	77.1	474	2 T20108	hypothetical prote
75	37	77.1	487	2 S42442	nuclear protein EB
76	37	77.1	547	2 C96828	unknown protein F1
77	37	77.1	667	2 T17231	hypothetical prote
78	37	77.1	695	2 T24950	hypothetical prote
79	37	77.1	745	2 T51370	hypothetical prote
80	37	77.1	876	2 T49801	hypothetical prote
81	37	77.1	1140	2 D88690	protein R4H10.3 {
82	37	77.1	1201	2 A57369	anillin - fruit fl
83	37	77.1	1262	2 T13353	protein stn-B - fr
84	37	77.1	1464	2 T13716	bazooka gene prote
85	37	77.1	1819	2 T32008	hypothetical prote
86	36.5	76.0	226	2 T39594	exeb protein - Aer
87	36.5	76.0	882	2 T38912	receptor tyrosine
88	36.5	76.0	890	1 A53743	ribosomal protein k
89	36	75.0	104	1 R5PW25	hypothetical prote
90	36	75.0	121	2 T04536	conserved hypothet
91	36	75.0	123	2 AH2707	hypothetical prote
92	36	75.0	169	2 H72470	hypothetical prote
93	36	75.0	196	2 T00702	hypothetical prote
94	36	75.0	199	2 T40919	probable signal re
95	36	75.0	217	2 S09754	hypothetical prote
96	36	75.0	224	2 D72861	Acorf-91 protein -
97	36	75.0	231	2 F82577	tonB protein XF228
98	36	75.0	235	2 G83625	hypothetical prote
99	36	75.0	238	2 T32889	hypothetical prote
100	36	75.0	256	2 A53340	H+-transporting tw
101	36	75.0	265	2 H75560	oxidoreductase, sh
102	36	75.0	267	2 JQ1752	hypothetical 30.6K

103	36	75.0	271	2	D70915	hypothetical prote	176	36	75.0	3175	1	RRWVEV	genome polyprotein
104	36	75.0	276	2	A12292	hypothetical prote	177	35	72.9	74	2	F36800	hypothetical prote
105	36	75.0	277	2	I38857	microtubule-associ	178	35	72.9	79	2	AE2788	hypothetical prote
106	36	75.0	285	2	T34615	NADH2 dehydrogenas	179	35	72.9	117	2	F72669	hypothetical prote
107	36	75.0	291	2	G84494	hypothetical prote	180	35	72.9	134	2	A72730	hypothetical prote
108	36	75.0	319	2	B87696	hypothetical prote	181	35	72.9	150	2	E72587	hypothetical prote
109	36	75.0	341	2	B69463	type I restriction	182	35	72.9	152	2	D95961	hypothetical prote
110	36	75.0	346	2	S76923	hypothetical prote	183	35	72.9	152	2	C72662	hypothetical prote
111	36	75.0	352	2	T18794	hypothetical prote	184	35	72.9	154	2	T41831	ACMNPV orf91 - Bom
112	36	75.0	361	2	S19552	potassium channel	185	35	72.9	154	2	E72624	hypothetical prote
113	36	75.0	378	2	T28112	hypothetical prote	186	35	72.9	179	2	F83305	hypothetical prote
114	36	75.0	380	2	T09661	ascorbate oxidase	187	35	72.9	195	2	H75455	conserved hypochet
115	36	75.0	393	2	T32533	hypothetical prote	188	35	72.9	197	2	T33525	hypothetical prote
116	36	75.0	417	2	G64417	hypothetical prote	189	35	72.9	198	2	T35752	hypothetical prote
117	36	75.0	426	2	F95058	hypothetical prote	190	35	72.9	201	2	D70864	hypothetical prote
118	36	75.0	438	2	S76184	hypothetical prote	191	35	72.9	203	2	T47571	hypothetical prote
119	36	75.0	431	2	S20055	nuclear factor I-X	192	35	72.9	209	2	A48232	cysteine-rich exte
120	36	75.0	435	2	A42672	choline kinase (EC	193	35	72.9	216	2	B37332	transforming prote
121	36	75.0	441	2	S37881	hypothetical prote	194	35	72.9	221	2	T15845	hypothetical prote
122	36	75.0	443	2	A38219	GAP-associated tyr	195	35	72.9	232	2	S24390	transforming prote
123	36	75.0	453	2	AX3342	choline kinase (EC	196	35	72.9	233	2	A37332	transforming prote
124	36	75.0	485	2	A33647	sulfated surface g	197	35	72.9	239	2	T36995	hypothetical prote
125	36	75.0	522	2	H97927	type I site-specif	198	35	72.9	255	2	AH0321	conserved hypochet
126	36	75.0	524	2	A75588	probable protein k	199	35	72.9	257	2	S32101	PHLPSA protein - c
127	36	75.0	577	2	T09034	proline-rich prote	200	35	72.9	262	2	B72037	ct598 hypothetical
128	36	75.0	598	2	T02265	hypothetical prote	201	35	72.9	262	2	E86588	ct598 hypothetical
129	36	75.0	604	2	H81110	sulfite reductase	202	35	72.9	262	2	A81504	conserved hypochet
130	36	75.0	606	1	UZAD12	terminal protein p	203	35	72.9	264	2	PQ0478	pistil extensin-li
131	36	75.0	609	2	T28736	hypothetical prote	204	35	72.9	285	1	B45393	polypeptide VI pre
132	36	75.0	618	2	T05518	hypothetical prote	205	35	72.9	268	2	T30630	hypothetical prote
133	36	75.0	645	2	T16078	hypothetical prote	206	35	72.9	269	2	T30468	hypothetical prote
134	36	75.0	649	2	AE1033	hypothetical prote	207	35	72.9	279	2	D96792	probable RNA-bindi
135	36	75.0	653	1	UZADP2	terminal protein p	208	35	72.9	293	2	C83845	superoxide dismuta
136	36	75.0	653	1	UZADP5	terminal protein p	209	35	72.9	309	2	G87498	hypothetical prote
137	36	75.0	679	2	A42073	potassium channel	210	35	72.9	313	1	FOVDA	gag polyprotein -
138	36	75.0	679	2	B75262	conserved hypochet	211	35	72.9	316	2	T31880	hypothetical prote
139	36	75.0	691	2	F91251	probable tape meas	212	35	72.9	321	2	F69451	cofactor modifying
140	36	75.0	715	1	TNBE77	77K alpha trans-in	213	35	72.9	358	2	A75561	conserved hypochet
141	36	75.0	745	2	T39952	hypothetical prote	214	35	72.9	368	2	T06093	hypothetical prote
142	36	75.0	769	2	I56546	Shaw type potassiu	215	35	72.9	369	2	B70968	hypothetical prote
143	36	75.0	772	2	T13078	KIAA0992 isoamylas	216	35	72.9	380	2	T32944	hypothetical prote
144	36	75.0	818	2	T02231	probable isomylas	217	35	72.9	384	2	H88924	hypothetical prote
145	36	75.0	822	2	T09221	exocyst complex pr	218	35	72.9	387	2	T21370	hypothetical prote
146	36	75.0	833	2	T01547	probable phospholi	219	35	72.9	389	2	S27200	proline-rich prote
147	36	75.0	885	2	S67660	hypothetical prote	220	35	72.9	392	2	JC2206	sensory kinase - S
148	36	75.0	927	2	T24031	hypothetical prote	221	35	72.9	393	2	PQ0479	pistil extensin-li
149	36	75.0	988	2	F86316	protein T1022.13	222	35	72.9	396	2	T04100	glucose-6-phosphat
150	36	75.0	992	1	GNWRK3	structural polypro	223	35	72.9	406	2	T44819	dolichyl-l-phosphate
151	36	75.0	1013	2	JC2314	chitin synthase (E	224	35	72.9	415	1	A34170	acrosin (EC 3.4.21
152	36	75.0	1014	2	JE0333	klotio protein - r	225	35	72.9	421	1	S11674	acrosin (EC 3.4.21
153	36	75.0	1048	2	T31425	C-terminal domain-	226	35	72.9	426	2	JQ1696	pistil extensin-li
154	36	75.0	1102	2	JC6316	probable protein k	227	35	72.9	436	2	T01652	zinc finger protei
155	36	75.0	1173	2	T31421	C-terminal domain-	228	35	72.9	439	2	S51939	chitinase (EC 3.2.
156	36	75.0	1200	2	T48194	hypothetical prote	229	35	72.9	440	2	T41766	ARIF-1 orf20/21 -
157	36	75.0	1239	2	G71266	probable Arp-depen	230	35	72.9	445	2	H96560	hypothetical prote
158	36	75.0	1262	2	T25168	hypothetical prote	231	35	72.9	451	2	T30603	perlecan homolog 2
159	36	75.0	1264	2	A36858	G2R protein - vari	232	35	72.9	456	2	S23104	choline kinase - h
160	36	75.0	1334	2	T50568	probable multi-dom	233	35	72.9	457	2	T50233	probable DNA-direc
161	36	75.0	1392	2	T51947	probable transcrip	234	35	72.9	471	2	E83592	hypothetical prote
162	36	75.0	1400	2	T52359	hypothetical prote	235	35	72.9	490	2	A35312	potassium channel
163	36	75.0	1422	2	T24212	hypothetical prote	236	35	72.9	523	2	A31948	carnitine octanoyl
164	36	75.0	1445	2	A59437	KIAA1204 protein l	237	35	72.9	530	2	A45690	transactivator EBN
165	36	75.0	1466	2	A36426	SPA2 protein - Yea	238	35	72.9	530	2	G70904	hypothetical prote
166	36	75.0	1557	2	T13160	protein CNK - frui	239	35	72.9	543	2	D83262	hypothetical prote
167	36	75.0	1575	2	T18545	lysobactin synthet	240	35	72.9	550	2	T36746	probable serine/th
168	36	75.0	1742	2	T49451	kinesin-like prote	241	35	72.9	563	2	A87378	hypothetical prote
169	36	75.0	1874	1	JQ0533	genome polyprotein	242	35	72.9	586	2	T29695	hypothetical prote
170	36	75.0	1896	2	B72175	Di5R protein - var	243	35	72.9	589	2	T05074	hypothetical prote
171	36	75.0	1897	2	T28621	hypothetical prote	244	35	72.9	598	2	S66669	potassium channel
172	36	75.0	2105	2	T18968	probable serine-ty	245	35	72.9	599	2	T10798	phosphorin-S - Vo
173	36	75.0	2133	2	T30637	hypothetical prote	246	35	72.9	602	2	JH0166	potassium voltage-
174	36	75.0	2459	2	AF2136	peptide synthetase	247	35	72.9	602	2	F84432	probable C2H2-type
175	36	75.0	2774	2	A43359	microtubule-associ	248	35	72.9	612	2	S60025	carnitine octanoyl

249	35	72.9	641	2	A54407	alpha-mannosidase	322	34	70.8	226	2	S41032	hypothetical prote
250	35	72.9	645	2	G01205	TYL protein - huma	323	34	70.8	226	2	C45343	hypothetical prote
251	35	72.9	650	2	B87791	protein B0207.1 [i	324	34	70.8	226	2	T42558	tegument protein 1
252	35	72.9	664	2	T01368	hypothetical prote	325	34	70.8	229	2	F84370	hypothetical prote
253	35	72.9	678	2	T04832	probable serine/th	326	34	70.8	238	2	T05263	hypothetical prote
254	35	72.9	681	2	JC5929	serine/arginine-ri	327	34	70.8	241	2	T22216	hypothetical prote
255	35	72.9	706	2	A45990	junctional sarcopl	328	34	70.8	245	2	T20831	hypothetical prote
256	35	72.9	708	2	D96711	hypothetical prote	329	34	70.8	247	2	G75356	hypothetical membr
257	35	72.9	734	2	T04876	hypothetical prote	330	34	70.8	248	2	F83558	hypothetical prote
258	35	72.9	749	2	S61643	probable membrane	331	34	70.8	251	2	F83101	conserved hypothet
259	35	72.9	774	1	JQ0550	1,4-alpha-glucan b	332	34	70.8	257	2	C72651	hypothetical prote
260	35	72.9	779	2	AC2249	heterocyst differe	333	34	70.8	275	2	G95962	probable taurine u
261	35	72.9	798	2	T21369	hypothetical prote	334	34	70.8	282	2	T26635	hypothetical prote
262	35	72.9	823	2	A36378	probable transcrip	335	34	70.8	283	2	S13383	hydroxyproline-ric
263	35	72.9	847	2	F96531	hypothetical prote	336	34	70.8	289	2	C96752	probable DNA-bindi
264	35	72.9	868	2	T20239	hypothetical prote	337	34	70.8	297	2	G70708	probable purC prot
265	35	72.9	899	2	F35299	hypothetical prote	338	34	70.8	298	2	H87533	peptidase, M23/M37
266	35	72.9	915	2	T12526	hypothetical prote	339	34	70.8	301	2	JQ1663	hybrid proline-ric
267	35	72.9	939	2	S28394	probable serine/th	340	34	70.8	301	2	T18788	hypothetical prote
268	35	72.9	944	2	T47246	chitin synthase (E	341	34	70.8	308	2	AH2896	transcription regu
269	35	72.9	974	2	E59434	Rho GTPase activat	342	34	70.8	308	2	B97672	rok family protein
270	35	72.9	1013	2	T33470	hypothetical prote	343	34	70.8	311	2	H70911	hypothetical prote
271	35	72.9	1041	2	T31097	chitin synthase (E	344	34	70.8	312	2	S04281	psbA intron 2 prot
272	35	72.9	1043	2	T19734	hypothetical prote	345	34	70.8	315	2	T06806	proline rich prote
273	35	72.9	1050	2	T27753	hypothetical prote	346	34	70.8	322	2	T23891	hypothetical prote
274	35	72.9	1175	2	S51005	protein-tyrosine-p	347	34	70.8	330	2	T05717	probable extensin
275	35	72.9	1176	2	I58345	protein-tyrosine p	348	34	70.8	335	2	S64352	hypothetical prote
276	35	72.9	1183	2	S63046	probable membrane	349	34	70.8	340	2	T36875	probable partition
277	35	72.9	1187	1	JC4155	protein-tyrosine-p	350	34	70.8	348	2	AD2247	hypothetical prote
278	35	72.9	1189	1	JC2366	protein-tyrosine-p	351	34	70.8	349	2	T47494	hypothetical prote
279	35	72.9	1245	1	VHWVB	structural polypro	352	34	70.8	353	2	B36963	bcsA 5'-region pro
280	35	72.9	1245	1	VHWVB2	structural polypro	353	34	70.8	358	2	B86141	protein T25K16.3 [
281	35	72.9	1249	2	T14150	vesicle associated	354	34	70.8	359	2	T13478	hypothetical prote
282	35	72.9	1258	2	JC5765	inositol polyphosp	355	34	70.8	361	2	T26610	hypothetical prote
283	35	72.9	1353	2	T00249	ichi protein - ink	356	34	70.8	369	2	AG1950	hypothetical prote
284	35	72.9	1428	2	T08852	lustrin A - Califo	357	34	70.8	372	1	T04157	dihydrokaempferol
285	35	72.9	1440	2	T27942	lin-15B protein -	358	34	70.8	380	2	T00683	hypothetical prote
286	35	72.9	1602	2	T31671	Rab3 GDP/GTP exch	359	34	70.8	386	2	E72645	hypothetical prote
287	35	72.9	1657	2	T19536	hypothetical prote	360	34	70.8	396	2	A58938	surface protein rh
288	35	72.9	1662	1	H71402	probable kinesin -	361	34	70.8	398	1	S41224	hnRNP protein - Af
289	35	72.9	1870	2	S37671	MHC class III hist	362	34	70.8	398	2	T37083	probable transposa
290	35	72.9	1872	2	S36152	MHC class III hist	363	34	70.8	407	2	S77922	shaggy-like protei
291	35	72.9	1879	2	T15962	hypothetical prote	364	34	70.8	409	2	D42971	tryparedoxin reduct
292	35	72.9	2186	2	T13169	tiggrin - fruit fl	365	34	70.8	409	2	S51105	shaggy protein kin
293	35	72.9	2241	2	S09811	hypothetical prote	366	34	70.8	410	2	C70889	hypothetical prote
294	35	72.9	2453	2	S60254	nuclear receptor c	367	34	70.8	412	2	S71266	shaggy-like protei
295	35	72.9	3119	2	I49729	HD protein - mouse	368	34	70.8	412	2	A84715	probable shaggy-li
296	34	71.9	59	2	A36589	bactenecin 7 - bov	369	34	70.8	418	2	F82638	hypothetical prote
297	34	70.8	81	2	B70727	hypothetical prote	370	34	70.8	420	2	A96613	probable glycogen
298	34	70.8	101	2	S29093	SMD3 protein - yea	371	34	70.8	420	2	T36193	probable salicylat
299	34	70.8	101	2	H72685	hypothetical prote	372	34	70.8	421	2	S51938	protein kinase AtK
300	34	70.8	106	2	T47814	hypothetical prote	373	34	70.8	423	2	T24383	hypothetical prote
301	34	70.8	111	2	T00688	hypothetical prote	374	34	70.8	432	2	AB2222	twitching motility
302	34	70.8	117	2	A33319	hypothetical prote	375	34	70.8	435	2	T46443	hypothetical prote
303	34	70.8	130	2	E72697	hypothetical prote	376	34	70.8	447	2	F86232	hypothetical prote
304	34	70.8	135	2	PC4252	hypothetical 135 p	377	34	70.8	447	2	T20249	hypothetical prote
305	34	70.8	141	2	JQ1669	hypothetical 16.1K	378	34	70.8	451	2	AI3623	acriflavin resista
306	34	70.8	144	2	H72464	hypothetical prote	379	34	70.8	452	2	D75572	conserved hypothet
307	34	70.8	145	2	T48552	glutaredoxin-like	380	34	70.8	463	2	S41495	dC stretch-binding
308	34	70.8	150	2	B27005	hypothetical prote	381	34	70.8	464	2	A54143	kappa-B motif-bind
309	34	70.8	152	2	JQ1244	coat protein 2 - b	382	34	70.8	464	2	S43363	transformation upr
310	34	70.8	152	2	T00877	hypothetical prote	383	34	70.8	464	2	A47655	apiclosome-associ
311	34	70.8	160	2	D48232	cysteine-rich exte	384	34	70.8	469	1	T02254	shaggy protein kin
312	34	70.8	161	2	E48232	cysteine-rich exte	385	34	70.8	469	1	T02256	shaggy protein kin
313	34	70.8	165	2	C48232	cysteine-rich exte	386	34	70.8	471	1	T02297	shaggy protein kin
314	34	70.8	175	1	GNVQFL	genome-linked prot	387	34	70.8	471	2	T03601	shaggy protein kin
315	34	70.8	175	2	D70824	hypothetical prote	388	34	70.8	474	2	S57239	gene forked protei
316	34	70.8	190	2	T35570	hypothetical prote	389	34	70.8	476	2	S57963	methyl CpG binding
317	34	70.8	198	2	D70509	hypothetical prote	390	34	70.8	477	2	I38409	adenyl cyclase-a
318	34	70.8	207	2	I53154	scleraxis - mouse	391	34	70.8	477	2	AE0608	conserved hypothet
319	34	70.8	211	2	B89716	protein F45B8.3 [i	392	34	70.8	478	2	F96594	unknown protein, 5
320	34	70.8	212	2	S74288	hypothetical prote	393	34	70.8	479	1	LYBELF	alpha trans-induci
321	34	70.8	219	2	F87505	hypothetical prote	394	34	70.8	483	1	TVRTKA	protein kinase fEC

395	34	70.8	485	2	PC4437	Nck, Ash and phosph	468	34	70.8	1429	2	T13720	gene expanded prot
396	34	70.8	490	1	IXBE33	alpha trans-induci	469	34	70.8	1436	2	S57238	forked protein 5.4
397	34	70.8	490	1	IXBE17	alpha trans-induci	470	34	70.8	1439	2	S57237	forked protein 5.6
398	34	70.8	490	2	JS0669	virion protein 16	471	34	70.8	1459	2	T30196	kinesin motor prot
399	34	70.8	491	2	T07598	proline-rich prote	472	34	70.8	1675	2	T31473	hypothetical prote
400	34	70.8	491	2	AG1838	hypothetical prote	473	34	70.8	1710	2	T14005	phospholipase D (G
401	34	70.8	492	2	C96521	protein F21D18.18	474	34	70.8	1776	2	G86280	protein T5E21.13 [
402	34	70.8	498	2	AC0442	ABC transporter AT	475	34	70.8	1872	2	T24683	hypothetical prote
403	34	70.8	499	2	S09880	hypothetical prote	476	34	70.8	2123	2	F86348	hypothetical prote
404	34	70.8	502	2	A55197	Wiskott-Aldrich sy	477	34	70.8	2212	2	T28157	erythrocyte membra
405	34	70.8	508	2	T45867	hypothetical prote	478	34	70.8	2649	2	T51023	hypothetical prote
406	34	70.8	520	1	QCQV65	hypothetical prote	479	34	70.8	7962	2	T38346	elastic titin - hu
407	34	70.8	529	2	T00677	hypothetical prote	480	33.5	69.8	1174	2	T49868	related to suppres
408	34	70.8	534	2	A37483	F protein - Muraya	481	33	68.8	52	2	E98316	hypothetical protei
409	34	70.8	542	2	A82965	hypothetical prote	482	33	68.8	55	2	S29770	DNA-binding protei
410	34	70.8	542	2	A44358	zyxin - chicken	483	33	68.8	59	2	D72852	actin rearrangemen
411	34	70.8	555	2	T30349	structural protein	484	33	68.8	70	2	A96850	protein F240I.6 [i
412	34	70.8	558	1	S68203	tyrosine N-monooxy	485	33	68.8	75	2	T26916	hypothetical prote
413	34	70.8	568	2	T40656	probable oxaly1-Co	486	33	68.8	76	2	E97763	protein transport
414	34	70.8	579	2	E75275	hypothetical prote	487	33	68.8	89	1	NSBOH7	nonhistone chromos
415	34	70.8	585	2	H86914	conserved hypotet	488	33	68.8	89	2	S11349	nonhistone chromos
416	34	70.8	587	2	T41653	probable transcrip	489	33	68.8	89	2	S33866	nonhistone chromos
417	34	70.8	589	2	AC1446	probable portal pr	490	33	68.8	90	1	NSCHH7	nonhistone chromos
418	34	70.8	589	2	S60154	serine/threonine-s	491	33	68.8	90	2	S03700	nonhistone chromos
419	34	70.8	589	2	S60159	serine/threonine-s	492	33	68.8	90	2	S01946	nonhistone chromos
420	34	70.8	604	2	E81905	probable sulfite r	493	33	68.8	105	1	NSCHH4	nonhistone chromos
421	34	70.8	604	2	S39885	forked protein - f	494	33	68.8	115	2	T36886	hypothetical prote
422	34	70.8	606	2	S70829	primase - Myxococ	495	33	68.8	125	2	A53892	synapsin I - mouse
423	34	70.8	616	2	AB1973	serine/threonine k	496	33	68.8	126	2	JU0068	homeotic protein C
424	34	70.8	621	2	B89451	probable zinc meta	497	33	68.8	132	2	T49536	hypothetical prote
425	34	70.8	626	1	NDBE61	exonuclease (EC 3	498	33	68.8	132	2	H96708	hypothetical prote
426	34	70.8	632	2	S38042	probable purine nu	499	33	68.8	139	2	T05847	hypothetical prote
427	34	70.8	635	2	H95918	probable DNA ligas	500	33	68.8	147	2	C75547	conserved hypotet
428	34	70.8	637	2	T04552	hypothetical prote	501	33	68.8	172	2	D41132	collagen-related p
429	34	70.8	641	2	S57236	forked protein 2.5	502	33	68.8	175	2	D75514	hypothetical prote
430	34	70.8	649	2	JN0809	drebrin E (clone g	503	33	68.8	187	2	C75558	acetyl-CoA carboxy
431	34	70.8	651	2	T14763	hypothetical prote	504	33	68.8	189	2	D49600	genome-linked prot
432	34	70.8	654	2	T34960	cell division prot	505	33	68.8	189	2	G70512	probable lpk prot
433	34	70.8	658	2	T04831	probable serine/th	506	33	68.8	196	2	B48232	cysteine-rich exte
434	34	70.8	662	2	D40228	neurexin II-beta p	507	33	68.8	199	2	AB3502	phosphatidyl-N-met
435	34	70.8	709	2	T46413	hypothetical prote	508	33	68.8	199	2	AB2171	hypothetical prote
436	34	70.8	726	2	G83310	conserved hypotet	509	33	68.8	206	2	T09503	probable heme tran
437	34	70.8	733	2	T12220	NADH2 dehydrogenas	510	33	68.8	207	2	C84631	modulin-like prote
438	34	70.8	736	2	I51691	dishevelled homolo	511	33	68.8	210	2	T50713	urease accessory p
439	34	70.8	741	2	D83633	conserved hypotet	512	33	68.8	212	2	E91046	transcription regu
440	34	70.8	751	2	T34490	hypothetical prote	513	33	68.8	212	2	F65023	gcvR protein - Esc
441	34	70.8	781	2	T02272	hypothetical prote	514	33	68.8	212	2	A85891	transcription regu
442	34	70.8	793	2	AH1094	probable peptidogl	515	33	68.8	214	2	S00763	H+-transporting tw
443	34	70.8	810	2	S65226	probable membrane	516	33	68.8	215	2	A69436	LSU ribosomal prot
444	34	70.8	814	2	T49207	receptor kinase-li	517	33	68.8	219	2	AH2953	conserved hypotet
445	34	70.8	817	2	S51342	varprolin - yeast	518	33	68.8	221	2	T07176	extensin homolog -
446	34	70.8	827	2	T04789	hypothetical prote	519	33	68.8	230	2	T29591	hypothetical prote
447	34	70.8	839	2	T04859	extensin homolog F	520	33	68.8	240	2	T25814	hypothetical prote
448	34	70.8	847	2	S75975	hypothetical prote	521	33	68.8	248	2	T40149	hypothetical prote
449	34	70.8	848	2	E71404	hypothetical prote	522	33	68.8	254	2	E98329	hypothetical prote
450	34	70.8	869	2	T22422	hypothetical prote	523	33	68.8	256	2	JQ1144	H+-transporting tw
451	34	70.8	872	2	T50369	probable serine/th	524	33	68.8	257	2	T29183	hypothetical prote
452	34	70.8	897	2	T02808	conserved hypotet	525	33	68.8	259	2	T22596	hypothetical prote
453	34	70.8	963	2	T48707	related to regulat	526	33	68.8	263	2	F70843	probable adhB prot
454	34	70.8	982	1	GNLJH2	pol polyprotein -	527	33	68.8	264	2	T06789	hydroxyproline-ric
455	34	70.8	986	2	T33135	hypothetical prote	528	33	68.8	265	2	AG2855	conserved hypotet
456	34	70.8	1038	2	S52522	hypothetical prote	529	33	68.8	265	2	D97632	hypothetical prote
457	34	70.8	1076	2	S50536	protein-tyrosine k	530	33	68.8	267	2	AC3133	proline rich prote
458	34	70.8	1091	2	S33596	leukocyte formin p	531	33	68.8	269	2	JC4899	conserved hypotet
459	34	70.8	1100	2	JC8033	hypothetical prote	532	33	68.8	272	2	AB2988	hypothetical prote
460	34	70.8	1111	2	T23047	protein-tyrosine-p	533	33	68.8	284	2	AB1919	collagen, cuticula
461	34	70.8	1174	2	T38140	brassinosteroid-in	534	33	68.8	286	2	S34565	hypothetical prote
462	34	70.8	1196	2	T09356	probable Ca2+-tran	535	33	68.8	288	2	T00473	hypothetical prote
463	34	70.8	1215	2	S50428	r19-1 protein - mo	536	33	68.8	293	2	S57376	hypothetical prote
464	34	70.8	1344	2	T14316	zinc finger protei	537	33	68.8	293	2	AE2018	hypothetical prote
465	34	70.8	1360	2	T06699	hypothetical prote	538	33	68.8	294	2	T34537	survival motor neu
466	34	70.8	1388	2	T00063	protein C47B8.8 [i	539	33	68.8	294	2	A55477	hypothetical prote
467	34	70.8	1425	2	B89303		540	33	68.8	294	2	A96155	

541	33	68.8	295	2	C95285	probable Lyg8-type	614	33	68.8	625	1	A43030	protein-tyrosine k
542	33	68.8	296	1	WNAD12	late 100K protein	615	33	68.8	627	2	D75393	serine proteinase,
543	33	68.8	299	2	G82616	tonB protein XF195	616	33	68.8	631	2	T32761	hypothetical prote
544	33	68.8	314	2	F96527	protein F27J15.20	617	33	68.8	633	2	S62057	proline-rich prote
545	33	68.8	315	2	F98295	hypothetical prote	618	33	68.8	636	2	H96666	probable phytochro
546	33	68.8	321	2	F84597	probable proline-r	619	33	68.8	638	1	QKMSPL	plasma kallikrein
547	33	68.8	325	2	S75095	hypothetical prote	620	33	68.8	647	2	T43952	hypothetical prote
548	33	68.8	343	2	H82171	probable multidrug	621	33	68.8	651	2	T31175	hypothetical prote
549	33	68.8	346	2	H84438	hypothetical prote	622	33	68.8	658	2	T08153	cysteine proteinas
550	33	68.8	347	1	SIQCQA	quinolinate synthe	623	33	68.8	659	2	A45184	protein-tyrosine k
551	33	68.8	347	2	AC0593	quinolinate synthe	624	33	68.8	659	2	T49553	protein-tyrosine k
552	33	68.8	347	2	C85577	quinolinate synthe	625	33	68.8	672	2	T32557	hypothetical prote
553	33	68.8	347	2	B99726	quinolinate synthe	626	33	68.8	682	1	S22700	amphiphysin - chic
554	33	68.8	360	1	A25732	inhibin alpha chai	627	33	68.8	685	2	T12139	NADH2 dehydrogenas
555	33	68.8	363	2	T16755	hypothetical prote	628	33	68.8	691	2	A25704	synapsin I - rat
556	33	68.8	364	2	F96603	hypothetical prote	629	33	68.8	695	2	T12148	NADH2 dehydrogenas
557	33	68.8	365	2	C37753	quinolinate syntha	630	33	68.8	695	2	S62400	amphiphysin (clone
558	33	68.8	372	2	T01600	hypothetical prote	631	33	68.8	704	2	A30411	synapsin Ia - rat
559	33	68.8	375	2	T48807	hypothetical prote	632	33	68.8	705	2	A35363	synapsin I splice
560	33	68.8	376	2	S71558	probable cell wall	633	33	68.8	706	2	E30411	synapsin Ia - bovi
561	33	68.8	376	2	B96698	unknown protein F1	634	33	68.8	707	2	A46302	PRB-associated spl
562	33	68.8	378	2	T46693	probable 2-nitropr	635	33	68.8	708	1	QQB88	Uu80 protein - hum
563	33	68.8	379	2	S31719	proline-rich prote	636	33	68.8	714	2	AD2236	p-aminobenzoic aci
564	33	68.8	383	2	AE2295	hypothetical prote	637	33	68.8	724	2	A38749	3-phosphatidylinos
565	33	68.8	384	2	T11580	probable 12-oxophy	638	33	68.8	728	2	T43632	serine/threonine p
566	33	68.8	387	2	T17682	hypothetical prote	639	33	68.8	731	2	T04455	hypothetical prote
567	33	68.8	389	2	B82819	outer membrane pro	640	33	68.8	733	1	S33643	transforming prote
568	33	68.8	390	2	A70656	hypothetical prote	641	33	68.8	736	2	T12217	NADH2 dehydrogenas
569	33	68.8	401	2	T51407	proline-rich prote	642	33	68.8	742	2	F84643	hypothetical prote
570	33	68.8	405	2	T37022	probable lycopene	643	33	68.8	747	1	E71429	probable FCA gamma
571	33	68.8	412	2	E88736	protein F3304.6a [644	33	68.8	759	2	T44142	DRI protein (impor
572	33	68.8	417	2	D70547	probable p1cA prot	645	33	68.8	765	2	S38099	hypothetical prote
573	33	68.8	419	2	F36788	hypothetical prote	646	33	68.8	767	2	I51083	SOX-L2 - rainbow t
574	33	68.8	419	2	T37577	hypothetical prote	647	33	68.8	768	2	D84271	small heat shock p
575	33	68.8	424	2	T33663	hypothetical prote	648	33	68.8	779	1	S61571	suppressor protein
576	33	68.8	427	2	E72488	probable tryptopha	649	33	68.8	792	2	T43630	serine/threonine p
577	33	68.8	428	2	F81694	pyruvate dehydroge	650	33	68.8	798	2	C96658	probable RNA helic
578	33	68.8	433	2	T02157	hypothetical prote	651	33	68.8	798	2	T19864	hypothetical prote
579	33	68.8	434	2	A75450	conserved hypoteth	652	33	68.8	802	2	T21315	hypothetical prote
580	33	68.8	437	2	A88942	protein R13D11.3 [653	33	68.8	807	1	WNAD15	late 100K protein
581	33	68.8	437	2	T32599	hypothetical prote	654	33	68.8	830	2	T10545	hypothetical prote
582	33	68.8	440	2	JC7807	Wiskott-Aldrich sy	655	33	68.8	837	2	T19825	hypothetical prote
583	33	68.8	448	2	A42026	cAMP response elem	656	33	68.8	845	2	E84488	probable RNA helic
584	33	68.8	448	2	T06076	proline-rich prote	657	33	68.8	884	2	T40690	hypothetical prote
585	33	68.8	456	2	B42036	cyclic AMP respons	658	33	68.8	891	2	H75507	hypothetical prote
586	33	68.8	461	2	G86498	Ct289 hypothetical	659	33	68.8	892	2	T09071	SH3 domains-contai
587	33	68.8	461	2	G72124	conserved hypoteth	660	33	68.8	900	2	A42024	transcription fact
588	33	68.8	464	2	C94686	probable AP2 domai	661	33	68.8	907	2	A24938	hypothetical T2 pr
589	33	68.8	464	2	S22697	extensin - Volvox	662	33	68.8	933	2	S17233	transcription fact
590	33	68.8	474	2	T20041	hypothetical prote	663	33	68.8	957	2	T03829	transcription fact
591	33	68.8	480	2	T10029	hypothetical prote	664	33	68.8	969	2	T02842	transable membrane
592	33	68.8	481	2	J80377	p70 S6 kinase (BC	665	33	68.8	985	2	T41135	hypothetical prote
593	33	68.8	481	2	T28900	hypothetical prote	666	33	68.8	986	2	T38205	RanBP7/importin-be
594	33	68.8	483	2	A12868	phosphomannomutase	667	33	68.8	998	2	T09492	general transcript
595	33	68.8	493	2	D83005	hypothetical prote	668	33	68.8	1003	2	T34066	hypothetical prote
596	33	68.8	494	2	AC3582	probable blue-copp	669	33	68.8	1005	2	T12546	hypothetical prote
597	33	68.8	496	2	T01564	hypothetical prote	670	33	68.8	1006	2	T42731	atrophin-1 related
598	33	68.8	505	1	S05380	transcription fact	671	33	68.8	1006	2	G86292	hypothetical prote
599	33	68.8	506	2	C97645	phosphomannomutase	672	33	68.8	1015	2	T42013	frequency clock pr
600	33	68.8	511	2	C70803	hypothetical prote	673	33	68.8	1016	2	H71460	probable outer mem
601	33	68.8	512	2	G86914	probable membrane	674	33	68.8	1075	2	T27623	hypothetical prote
602	33	68.8	512	2	B70786	probable integral	675	33	68.8	1080	2	T27622	hypothetical prote
603	33	68.8	524	2	T04564	cytochrome P450 ho	676	33	68.8	1122	2	T14180	exit protein - Myc
604	33	68.8	526	2	T04566	cytochrome P450 ho	677	33	68.8	1151	2	T18535	high molecular mas
605	33	68.8	531	2	B85049	probable transposo	678	33	68.8	1201	2	T29329	hypothetical prote
606	33	68.8	533	2	A85181	FCA delta protein	679	33	68.8	1210	2	T39410	AP-4 protein, spli
607	33	68.8	551	2	S57447	HPBRII-7 protein -	680	33	68.8	1213	2	A58198	serine/proline-ric
608	33	68.8	572	2	G02845	zyxin - human	681	33	68.8	1215	2	T32734	myosin-IA - Acanth
609	33	68.8	574	2	B35149	ipaH protein - Shi	682	33	68.8	1226	2	S69078	hypothetical prote
610	33	68.8	577	2	T28017	hypothetical prote	683	33	68.8	1403	2	S24548	homeotic protein p
611	33	68.8	592	2	T21536	hypothetical prote	684	33	68.8	1536	2	S59841	4-alpha-glucanotra
612	33	68.8	601	2	S56144	SH3 domain binding	685	33	68.8	1556	2	A60988	saliva-interacting
613	33	68.8	620	1	S33253	protein-tyrosine k	686	33	68.8	1561	1	S06839	surface antigen sp

687	33	68.8	1565	2	S04729	surface antigen pa	760	32	66.7	233	2	H75359	hypothetical prote
688	33	68.8	1613	2	JE0272	low density lipopr	761	32	66.7	235	1	RWHUT8	T-cell surface gly
689	33	68.8	1613	2	JE0273	low density lipopr	762	32	66.7	235	2	F83448	succinate dehydrog
690	33	68.8	1696	2	T00057	hypothetical prote	763	32	66.7	235	2	G81138	probable succinate
691	33	68.8	1753	2	S30855	hypothetical prote	764	32	66.7	235	2	A83970	hypothetical prote
692	33	68.8	1872	2	T00339	hypothetical prote	765	32	66.7	236	2	F82118	succinate dehydrog
693	33	68.8	1888	2	T14273	zinc finger protei	766	32	66.7	237	2	A47368	HNF-3/fork head tr
694	33	68.8	2222	2	A37490	voltage-dependent	767	32	66.7	238	1	DEEC81	succinate dehydrog
695	33	68.8	2251	2	S54972	voltage-dependent	768	32	66.7	238	2	B90722	succinate dehydrog
696	33	68.8	2293	2	S29236	calcium channel pr	769	32	66.7	238	2	D85573	succinate dehydrog
697	33	68.8	2270	2	A54972	voltage-dependent	770	32	66.7	238	2	AH0136	succinate dehydrog
698	33	68.8	2272	2	C54972	voltage-dependent	771	32	66.7	239	2	AC0591	probable prolinae-r
699	33	68.8	2764	2	T13949	neurofibromin - fr	772	32	66.7	239	2	F84749	nitrate transport
700	33	68.8	2802	2	T13945	neurofibromin - fr	773	32	66.7	248	2	G69329	hypothetical prote
701	33	68.8	2802	2	T13947	neurofibromin - fr	774	32	66.7	248	2	S75632	nitrate transport
702	33	68.8	3010	1	GNWVTC	genome polyprotein	775	32	66.7	248	2	F91197	Espp protein [impo
703	33	68.8	3716	2	E70969	probable PPE prote	776	32	66.7	248	2	B86044	Espp [imported] -
704	33	68.8	4861	2	S71752	giant protein p619	777	32	66.7	255	2	JC7264	CCAAAT-enhancer bin
705	33	68.8	5126	2	S40450	ryanodine receptor	778	32	66.7	257	2	F81096	conserved hypoteth
706	32.5	67.7	353	2	G02729	thrombopoietin - h	779	32	66.7	267	1	B64351	nitrate transport
707	32.5	67.7	353	2	I80105	thrombopoietin pre	780	32	66.7	268	2	H81844	conserved hypoteth
708	32.5	67.7	623	2	C84694	probable protein k	781	32	66.7	268	2	A49303	homeotic protein C
709	32.5	67.7	687	2	T08877	Modin - Podospora	782	32	66.7	270	2	E87649	hypothetical prote
710	32.5	67.7	1089	2	T14576	nosa protein - sli	783	32	66.7	271	2	S72782	Bl496 F2.81 protei
711	32.5	67.7	1456	2	T01397	LTR gag/pol polypr	784	32	66.7	272	2	D83154	probable permease
712	32.5	67.7	2706	2	T28155	variant-specific s	785	32	66.7	272	2	B75414	molybdenum cofacto
713	32	66.7	51	2	S64676	acetylcholinestera	786	32	66.7	276	2	AB3644	hypothetical cyto
714	32	66.7	80	2	I65235	testicular luteini	787	32	66.7	278	2	B75572	conserved hypoteth
715	32	66.7	101	2	S04804	hypothetical prote	788	32	66.7	278	2	A49266	fas ligand - rat
716	32	66.7	101	2	C72540	hypothetical prote	789	32	66.7	278	2	AB2254	hypothetical prote
717	32	66.7	108	2	B95314	hypothetical prote	790	32	66.7	279	2	T10361	probable oligopept
718	32	66.7	118	2	PN0141	lutropin beta chai	791	32	66.7	283	2	D71130	hypothetical prote
719	32	66.7	118	2	PN0139	lutropin beta chai	792	32	66.7	284	2	B70834	hypothetical prote
720	32	66.7	119	2	A61465	lutropin beta chai	793	32	66.7	284	2	D81075	conserved hypoteth
721	32	66.7	119	2	D72520	hypothetical prote	794	32	66.7	284	2	D81866	hypothetical prote
722	32	66.7	128	2	T05635	hypothetical prote	795	32	66.7	284	2	G96815	probable secreted
723	32	66.7	130	1	ERAD53	early E3B 14.5K pr	796	32	66.7	286	2	F87220	aminomethyltransfe
724	32	66.7	133	2	JC2202	secretin precursor	797	32	66.7	287	2	AD3412	aminoglycoside 6-a
725	32	66.7	134	2	D84672	hypothetical prote	798	32	66.7	292	2	B83690	conserved hypoteth
726	32	66.7	135	2	H72639	hypothetical prote	799	32	66.7	298	2	D95314	transcription acti
727	32	66.7	138	2	S00512	lutropin beta chai	800	32	66.7	299	2	AD0058	hypothetical prote
728	32	66.7	139	2	I52320	testicular luteini	801	32	66.7	299	2	T17832	hypothetical prote
729	32	66.7	141	1	UTPOB	lutropin beta chai	802	32	66.7	301	2	T33702	hypothetical prote
730	32	66.7	141	1	UTRFB	lutropin beta chai	803	32	66.7	304	2	T16704	hypothetical prote
731	32	66.7	141	2	JC4527	lutinizing hormon	804	32	66.7	306	2	T06435	ribonuclease S5 ho
732	32	66.7	142	2	S11438	coat protein 2 - b	805	32	66.7	306	2	T52340	cell wall-plasma m
733	32	66.7	143	1	A60159	interleukin-3 prec	806	32	66.7	315	2	H83276	probable lipase PA
734	32	66.7	145	2	F70919	hypothetical prote	807	32	66.7	324	2	T22580	hydroxyproline-ric
735	32	66.7	147	2	S37485	gene msl1 protein	808	32	66.7	328	2	JQ0985	probable oxidoredu
736	32	66.7	148	2	S39206	proline-rich prote	809	32	66.7	334	2	E70519	chitinase (EC 3.2.
737	32	66.7	150	2	C72697	hypothetical prote	810	32	66.7	334	2	S20982	probable histone p
738	32	66.7	153	2	T14303	heat shock protein	811	32	66.7	338	2	T40774	structural protein
739	32	66.7	154	2	C71162	hypothetical prote	812	32	66.7	340	2	S32584	unknown protein [i
740	32	66.7	157	2	S35037	helix-loop-helix p	813	32	66.7	342	2	A96511	rod shape-determin
741	32	66.7	158	1	ZBBE11	17K zinc-binding p	814	32	66.7	346	2	T51252	rod shape-determin
742	32	66.7	161	2	H85120	probable prolinae-r	815	32	66.7	346	2	G87440	hypothetical prote
743	32	66.7	166	2	T09593	CIC protein, cold-	816	32	66.7	346	2	T27896	hydroxyproline-ric
744	32	66.7	183	2	I38923	microfibril-associ	817	32	66.7	350	2	S22456	extensin homolog T
745	32	66.7	183	2	A49313	microfibril-associ	818	32	66.7	350	2	T14191	MutT/nudix family
746	32	66.7	185	1	JC2394	phospholipase A2 1	819	32	66.7	350	2	G75571	quinolinate synth
747	32	66.7	185	2	D75602	hypothetical prote	820	32	66.7	353	2	E82151	prostanoid dp rece
748	32	66.7	187	2	E95413	conserved hypoteth	821	32	66.7	359	2	I39153	hypothetical prote
749	32	66.7	192	2	F41816	amelogenin (Y chro	822	32	66.7	359	2	T22774	hydroxyproline-ric
750	32	66.7	198	2	T07780	remorin - potato	823	32	66.7	369	2	S20500	CT058 hypotethet
751	32	66.7	201	2	G70874	probable transcrip	824	32	66.7	371	2	H86536	conserved hypoteth
752	32	66.7	204	2	B72733	probable membrane	825	32	66.7	371	2	B72087	hypothetical prote
753	32	66.7	206	2	T29696	hypothetical prote	826	32	66.7	374	2	T08747	hypothetical prote
754	32	66.7	212	2	T47947	hypothetical prote	827	32	66.7	384	2	T23604	hypothetical prote
755	32	66.7	219	2	T10563	proline-rich prote	828	32	66.7	391	2	T36311	probable lipase -
756	32	66.7	219	2	E83299	hypothetical prote	829	32	66.7	398	1	WZBEN3	protein-serine/thr
757	32	66.7	220	2	S42879	modulin-30 - kidne	830	32	66.7	399	2	CT6522	paired-type homeod
758	32	66.7	226	2	S33024	hypothetical prote	831	32	66.7	403	2	S52796	prpL2 protein - hu
759	32	66.7	229	2	D95398	hypothetical prote	832	32	66.7	411	2	G86923	probable hydrolase

833	32	66.7	416	1	A31959	lysosome-associate	906	32	66.7	636	2	A12552	transposase alt-801
834	32	66.7	421	2	C82253	folypolyglutamate	907	32	66.7	638	2	A33991	somatotropin recep
835	32	66.7	424	2	S09884	hypothetical prote	908	32	66.7	638	2	S12136	somatotropin recep
836	32	66.7	429	2	H71539	probable dihydroli	909	32	66.7	638	2	B28176	somatotropin recep
837	32	66.7	433	2	T07910	hydroxyproline-ric	910	32	66.7	638	2	A33505	hypothetical prote
838	32	66.7	434	2	F75425	tRNA nucleotidyltr	911	32	66.7	642	2	T27429	hypothetical prote
839	32	66.7	435	2	JN0290	protein-tyrosine k	912	32	66.7	643	2	T27429	hypothetical prote
840	32	66.7	442	2	T16035	hypothetical prote	913	32	66.7	650	2	A03352	chaperone protein
841	32	66.7	443	2	T17220	hypothetical prote	914	32	66.7	651	2	JC7705	death receptor-6 -
842	32	66.7	444	2	E83802	hypothetical prote	915	32	66.7	651	2	F86563	hypothetical prote
843	32	66.7	446	2	A44418	H-2 region II bind	916	32	66.7	651	2	A39372	potassium channel
844	32	66.7	447	2	T49350	hypothetical prote	917	32	66.7	651	2	A72060	hypothetical prote
845	32	66.7	448	2	D41727	retinoid X recepto	918	32	66.7	656	2	S55262	UV-endonuclease -
846	32	66.7	449	2	S16748	proline-rich prote	919	32	66.7	657	2	S32739	filensin - chicken
847	32	66.7	451	2	A41651	retinoic acid rece	920	32	66.7	662	2	I37892	IL12 receptor comp
848	32	66.7	454	1	QMS71	tumor necrosis fac	921	32	66.7	662	2	T17339	hypothetical prote
849	32	66.7	458	1	KHR20A	oryzain (EC 3.4.22	922	32	66.7	663	2	T14093	hmf-3/forkhead tra
850	32	66.7	472	2	T24316	hypothetical prote	923	32	66.7	668	2	C71868	hypothetical prote
851	32	66.7	472	2	H70836	hypothetical prote	924	32	66.7	676	2	F69394	H+-transporting AT
852	32	66.7	475	2	J32036	hypothetical prote	925	32	66.7	677	2	S33608	somatotropin-bindi
853	32	66.7	475	2	S47861	hypothetical prote	926	32	66.7	680	2	JC5895	killer cell inhibi
854	32	66.7	477	1	I51645	mei-1 protein - Ca	927	32	66.7	683	2	T12295	NADH2 dehydrogenas
855	32	66.7	477	2	JN0291	protein-tyrosine k	928	32	66.7	685	2	T18964	hypothetical prote
856	32	66.7	482	2	A39285	calcitonin recepto	929	32	66.7	694	2	T12675	NADH2 dehydrogenas
857	32	66.7	485	2	F87235	conserved membrane	930	32	66.7	695	2	A87494	topoisomerase IV,
858	32	66.7	489	2	A39257	cytochrome P450 2C	931	32	66.7	715	2	B83518	pineloyl-CoA synth
859	32	66.7	490	2	S28166	cytochrome P450 2C	932	32	66.7	724	2	A38748	3--phosphatidylinos
860	32	66.7	490	2	A29782	unspecific monooxy	933	32	66.7	724	2	A83747	phosphatidylinoisit
861	32	66.7	490	2	I48189	cytochrome P450 II	934	32	66.7	731	2	AE1945	hypothetical prote
862	32	66.7	490	2	B28516	cytochrome P450 2C	935	32	66.7	738	2	I49295	IL-12 receptor bet
863	32	66.7	496	2	T09300	beta-amyase (EC 3	936	32	66.7	740	2	T12753	NADH2 dehydrogenas
864	32	66.7	496	2	T11575	beta-amyase (EC 3	937	32	66.7	747	2	T13683	NADH2 dehydrogenas
865	32	66.7	498	2	I47130	calcitonin recepto	938	32	66.7	758	2	T15577	hypothetical prote
866	32	66.7	503	2	S62018	probable membrane	939	32	66.7	760	2	T06291	extensin homolog I
867	32	66.7	506	2	AQ1864	hypothetical prote	940	32	66.7	763	2	JC5234	amine oxidase (cop
868	32	66.7	511	2	A40110	nicotinic acetylch	941	32	66.7	780	2	T00366	hypothetical prote
869	32	66.7	514	2	H89695	protein M0386.3 (i	942	32	66.7	782	2	T15060	hypothetical prote
870	32	66.7	516	2	H82756	conserved hypotet	943	32	66.7	785	2	S75819	general secretion
871	32	66.7	517	2	B89530	protein H28G03.2 (944	32	66.7	795	2	T49835	hypothetical prote
872	32	66.7	520	1	NNEB1	anthranilate synth	945	32	66.7	796	1	A32434	hypothetical prote
873	32	66.7	520	1	AF0653	anthranilate synth	946	32	66.7	796	2	E96654	hypothetical prote
874	32	66.7	520	2	I84718	RXR-beta1 isoform	947	32	66.7	797	2	T33098	probable b2ip prot
875	32	66.7	526	2	T33047	hypothetical prote	948	32	66.7	798	2	D96563	hypothetical prote
876	32	66.7	527	2	S13763	protein-tyrosine k	949	32	66.7	811	2	S08579	hypothetical prote
877	32	66.7	533	2	S37781	retinoid X recepto	950	32	66.7	821	2	T34232	hypothetical prote
878	32	66.7	534	2	S21961	proline-rich prote	951	32	66.7	830	2	I50455	prolactin receptor
879	32	66.7	545	2	JC7978	proctolin receptor	952	32	66.7	831	2	JQ1655	conserved hypotet
880	32	66.7	546	2	D83408	hypothetical prote	953	32	66.7	833	2	A90575	related to cytosin
881	32	66.7	548	2	T41182	probable protein p	954	32	66.7	845	2	T52518	hypothetical prote
882	32	66.7	549	2	A86183	hypothetical prote	955	32	66.7	845	2	T17291	testicular dynamin
883	32	66.7	559	2	I49444	SH3 binding protei	956	32	66.7	848	2	I55498	penicillin-binding
884	32	66.7	561	2	G96752	unknown protein F2	957	32	66.7	873	2	B75514	Cell division prot
885	32	66.7	570	1	Q0BE2R	BVRF1 (EC-RF2) pro	958	32	66.7	886	2	C87031	Cell division prot
886	32	66.7	579	2	T30635	hypothetical prote	959	32	66.7	892	2	T15195	hypothetical prote
887	32	66.7	581	2	I45971	prolactin receptor	960	32	66.7	903	1	VGEBE1	glycoprotein B pre
888	32	66.7	586	2	A41125	gamma-glutamyltran	961	32	66.7	928	2	T38419	Sec76 domain prote
889	32	66.7	592	2	T07083	hypothetical prote	962	32	66.7	945	2	T23965	hypothetical prote
890	32	66.7	593	2	S49525	glycoprotein G - s	963	32	66.7	948	2	A57640	retinoblastoma bin
891	32	66.7	594	2	D64676	oligopeptide ABC t	964	32	66.7	954	1	S68178	mixed-lineage prot
892	32	66.7	595	2	D71841	hypothetical prote	965	32	66.7	974	1	A40213	optic lobe develop
893	32	66.7	602	2	JU0215	protein-tyrosine k	966	32	66.7	978	2	T03763	BAP-135 protein ho
894	32	66.7	603	2	A75373	probable N-acetyl	967	32	66.7	980	2	S54986	regulatory protein
895	32	66.7	615	1	KFHU12	coagulation factor	968	32	66.7	990	1	TQEC26	transposase - Bach
896	32	66.7	620	2	S06733	hydroxyproline-ric	969	32	66.7	1003	2	C71139	hypothetical prote
897	32	66.7	624	2	JC5471	regulatory protein	970	32	66.7	1006	2	F87363	medium-chain-fatty
898	32	66.7	628	2	S01955	hypothetical prote	971	32	66.7	1023	2	E71376	conserved hypotet
899	32	66.7	630	1	T01380	protein-tyrosine k	972	32	66.7	1027	2	F82829	acriflavin resista
900	32	66.7	630	2	T02524	probable RING zinc	973	32	66.7	1039	2	T22117	hypothetical prote
901	32	66.7	630	2	I51086	prolactin receptor	974	32	66.7	1046	2	F71432	hypothetical prote
902	32	66.7	631	2	I56997	protein-tyrosine k	975	32	66.7	1070	2	T25591	hypothetical prote
903	32	66.7	631	2	T31782	hypothetical prote	976	32	66.7	1087	2	S58147	protein kinase - f
904	32	66.7	633	2	S47144	matting type A prot	977	32	66.7	1097	1	S47220	protein kinase C (
905	32	66.7	634	2	S33339	somatotropin recep	978	32	66.7	1102	2	T28666	protein kinase C-r

transcription fact
DNA-directed DNA p
genome polyprotein
hypothetical prote
hypothetical prote
hypothetical prote
protein-tyrosine k
T20H2.9 protein -
hypothetical prote
avrBs3-2 protein -
hypothetical prote
hypothetical 125K
hypothetical prote
KIAA0641 protein [i
RNA-directed DNA p
SWT4 related prote
myosin heavy chain
dynactin 1 - mouse
period protein hom
period protein hom
hypothetical prote

32 66.7 1104 2 I38869
32 66.7 1106 1 A39299
32 66.7 1107 1 JQ1658
32 66.7 1119 2 T16720
32 66.7 1122 2 T47424
32 66.7 1128 2 AH1949
32 66.7 1136 1 S57845
32 66.7 1137 2 A86335
32 66.7 1141 2 T29185
32 66.7 1160 2 S34809
32 66.7 1173 2 T25985
32 66.7 1187 2 JQ0316
32 66.7 1199 2 T15826
32 66.7 1201 2 G86441
32 66.7 1207 2 T00378
32 66.7 1212 2 S27771
32 66.7 1240 2 T48800
32 66.7 1271 2 S37958
32 66.7 1281 2 JC5368
32 66.7 1290 2 T00018
32 66.7 1291 2 T00019
32 66.7 1308 2 T15280

ALIGNMENTS

RESULT 1
A56031
potassium channel KCNA5 - human
N/Alternate names: potassium channel HK2; potassium channel PCN1; shaker-related potass
C/Species: Homo sapiens (man)
C/Date: 05-Apr-1995 #sequence revision 05-Apr-1995 #text_change 09-Jul-2004
C/Accession: A56031; A38556; B39922; A38074
R/Phillipson, L.H.; LaMendola, J.; Bell, G.I.; Steiner, D.F.
submitted to GenBank, September 1990
A/Reference number: A56031
A/Accession: A56031
A/Molecule type: mRNA
A/Residues: 1-613 <PHI>
A/Cross-references: UNIPROT:P22460; UNIPARC:UPI000016AE76; GB:M55513; NID:G189653; PIDN:
R/Phillipson, L.H.; Hice, R.E.; Schaefer, K.; LaMendola, J.; Bell, G.I.; Nelson, D.J.; St
Proc. Natl. Acad. Sci. U.S.A. 88, 53-57, 1991
A/Title: Sequence and functional expression in Xenopus oocytes of a human insulinoma and
A/Reference number: A38556; MUID:91095456; PMID:1986382
A/Accession: A38556
A/Status: nucleic acid sequence not shown
A/Molecule type: DNA
A/Residues: 1-56, 'G', 58-137, 'L', 139-213, 'R', 215-227, 'P', 229-545, 'QG', 546-613 <PH2>
A/Cross-references: UNIPARC:UPI00001779F3; GB:M55513
R/Tamkun, M.M.; Knoth, K.M.; Walbridge, J.A.; Kroemer, H.; Roden, D.M.; Glover, D.M.
FASEB J. 5, 331-337, 1991
A/Title: Molecular cloning and characterization of two voltage-gated K(+) channel cDNAs
A/Reference number: A39922; MUID:91160866; PMID:2001794
A/Accession: B39922
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-54, 56-137, 'L', 139-186, 'G', 189-213, 'R', 215-227, 'P', 229-297, 'PTQRGH', 309-558
A/Cross-references: UNIPARC:UPI00001779F4; GB:M60451
R/Curran, M.E.; Landes, G.M.; Keating, M.T.
Genomics 12, 729-737, 1992
A/Title: Molecular cloning, characterization, and genomic localization of a human potass
A/Reference number: A38074; MUID:92241872; PMID:1349297
A/Accession: A38074
A/Molecule type: DNA
A/Residues: 1-137, 'L', 139-153, 'R', 155-213, 'R', 215-227, 'P', 229-281, 'V', 283-578, 'QLPPREV'
A/Cross-references: UNIPARC:UPI000016AEEA; GB:M83254; NID:G190202; PIDN:AAA60146.1; PID:
A/Experimental source: heart
A/Note: sequence extracted from NCBI backbone (NCBIN:98573, NCBI:P:98577)
C/Genetics:
A/Gene: GDB:KCNA5
A/Cross-references: GDB:I27904; OMIM:I76267
A/Map position: 12p13.33-12p13.31

C/Superfamily: potassium channel protein drk1
C/Keywords: glycoprotein; phosphoprotein; potassium channel; transmembrane protein; volt
F:125,190/Binding site: carboxylate (Asn) (covalent) #status Predicted
F:557/Binding site: phosphate (Ser) (covalent) #status Predicted

Query Match 97.9%; Score 47; DB 2; Length 613;
Best Local Similarity 88.9%; Pred. No. 18;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 RPLPPLXP 12
Db 65 RPLPPLPDP 73
|||||
|

RESULT 2
T07173
hypothetical protein SEND32, senescence down-regulated - tomato (fragment)
C/Species: Lycopersicon esculentum (tomato)
C/Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 09-Jul-2004
C/Accession: T07173
R/John, I.; Hackett, R.; Cooper, W.; Drake, R.; Farrell, A.; Grierson, D.
Plant Mol. Biol. 33, 641-651, 1997
A/Title: Cloning and characterization of tomato leaf senescence-related cDNAs.
A/Reference number: Z15976; MUID:97238471; PMID:9132056
A/Accession: T07173
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-162 <JOH>
A/Cross-references: UNIPROT:Q43516; UNIPARC:UPI000009DEBF; EMBL:Z75519; NID:G1418979; PI
A/Experimental source: cultivar Ailsa Craig; leaf; clone SRND32
C/Superfamily: herpesvirus immediate-early protein IE110; RING finger homology

Query Match 89.6%; Score 43; DB 2; Length 162;
Best Local Similarity 87.5%; Pred. No. 16;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 PLPPLXP 12
Db 155 PLPPLPSP 162
|||||
|

RESULT 3
E89008
protein W08A12.3 [imported] - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C/Accession: E89008
R/Anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A/Title: Genome sequence of the nematode C. elegans: a platform for investigating biology
A/Reference number: A75000; MUID:99059613; PMID:9851916
A/Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_eleg
A/Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A/Accession: E89008
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-198 <STO>
A/Cross-references: UNIPROT:O01621; UNIPARC:UPI0000076A91; GB:chr_V; PIDN:AAB52319.1; PI
C/Genetics:
A/Gene: W08A12.3
A/Map position: 5

Query Match 87.5%; Score 42; DB 2; Length 198;
Best Local Similarity 87.5%; Pred. No. 28;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 PLPPLXP 12
Db 98 PLPPLPSP 105
|||||
|

RESULT 4
I38707

Fas ligand - human
C:Species: Homo sapiens (man)
C>Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 09-Jul-2004
C/Accession: I38707; JC2340; S57565; I38554
R;Takahashi, T.; Tanaka, M.; Inazawa, J.; Abe, T.; Suda, T.; Nagata, S.
Int. Immunol. 6, 1567-1574, 1994
A;Title: Human Fas ligand: gene structure, chromosomal location and species specificity.
A;Reference number: I38707; MUID:95127560; PMID:7826947
A;Accession: I38707
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-281 <RES>
A;Cross-references: UNIPROT:P48023; UNIPARC:UPI000000D91A; EMBL:U11821; NID:G595430; PID
R;Mita, E.; Hayaishi, N.; Iio, S.; Takehara, T.; Hijioaka, T.; Kasahara, A.; Fusamoto, H.;
Biochem. Biophys. Res. Commun. 204, 468-474, 1994
A;Title: Role of Fas ligand in apoptosis induced by hepatitis C virus infection.
A;Reference number: JC2340; MUID:95071350; PMID:7980502
A;Accession: JC2340
A;Molecule type: DNA
A;Residues: 1-281 <MIT>
A;Cross-references: UNIPARC:UPI000000D91A; GB:D38122; DBJ:D29820; NID:G601892; PIDN:BAR
R;Schatzlein, C.E.
submitted to the EMBL Data Library, June 1995
A;Reference number: S57565
A;Accession: S57565
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-281 <SCH>
A;Cross-references: UNIPARC:UPI000000D91A; EMBL:X89102; NID:G887455; PID:G887456
R;Alderson, M.R.; Tough, T.W.; Davis-Smith, T.; Braddy, S.; Falk, B.; Schooley, K.A.; Go
J. Exp. Med. 181, 71-77, 1995
A;Title: Fas ligand mediates activation-induced cell death in human T lymphocytes.
A;Reference number: I38554; MUID:95105731; PMID:7528780
A;Accession: I38554
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-281 <RE2>
A;Cross-references: UNIPARC:UPI000000D91A; EMBL:U08137; NID:G624627; PIDN:AAC50071.1; PI
C;Genetics:
A;Gene: FasL
A;Introns: 151/1; 116/3
C;Keywords: glycoprotein; transmembrane protein
F;80-102/Domain: transmembrane #status predicted <TMM>
F;76,184,250,260/Binding site: carbonylate (Asn) (covalent) #status predicted

Query Match 87.5%; Score 42; DB 2; Length 281;
Best Local Similarity 87.5%; Pred. No. 40;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 PLPPLPXP 12
DB 62 PLPPLPLP 69

RESULT 5
S50755
hypothetical protein VSP-3 - Chlamydomonas reinhardtii
C;Species: Chlamydomonas reinhardtii
C>Date: 14-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 09-Jul-2004
C/Accession: S50755
R;Moensner, J.P.; Molendijk, A.J.; van Egmond, P.; Klis, F.M.; Goodenough, U.W.; Haring,
Plant Mol. Biol. 26, 947-960, 1994
A;Title: Domain conservation in several volvocalean cell wall proteins.
A;Reference number: S50754; MUID:95093034; PMID:8000007
A;Accession: S50755
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-473 <WOE>
A;Cross-references: UNIPROT:Q39620; UNIPARC:UPI000000D9F5C; EMBL:L29029; NID:G530875; PID

Query Match 87.5%; Score 42; DB 2; Length 473;
Best Local Similarity 77.8%; Pred. No. 69;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 RPLPPLPXP 12
DB 360 RPSPLPLPSP 368

RESULT 6

JC7509
glycoprotein VI-1 - human
C;Species: Homo sapiens (man)
C>Date: 30-Jun-2001 #sequence_revision 30-Jun-2001 #text_change 09-Jul-2004
C/Accession: JC7509; PC7101
R;Ezumi, Y.; Uchiyama, T.; Takayama, H.
Biochem. Biophys. Res. Commun. 277, 27-36, 2000
A;Title: Molecular cloning, genomic structure, chromosomal localization, and alternative
A;Reference number: JC7509; MUID:20483673; PMID:11027634
A;Contents: Platelet
A;Accession: JC7509
A;Molecule type: mRNA
A;Residues: 1-339 <EZU>
A;Cross-references: UNIPROT:Q9UIF2; UNIPARC:UPI000006F4A8; DBJ:AB043819
A;Accession: PC7101
A;Molecule type: protein
A;Residues: 28-41; 62-79; 114-142 <EZ2>
A;Cross-references: UNIPARC:UPI000017A509; UNIPARC:UPI000017A50A; UNIPARC:UPI000017A50B
C;Comment: This protein, which belongs to the immunoglobulin superfamily, is the major c
or gamma chain as a signal transducing subunit, and plays some roles in cancer cells.
C;Genetics:
A;Gene: gpVI-1
A;Map position: 19q13.4
A;Introns: 62/1; 95/1; 353/1; 638/1; 692/1; 752/1; 803/1
C;Keywords: glycoprotein; immunoglobulin; platelet

Query Match 85.4%; Score 41; DB 2; Length 339;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 RPLPPLP 10
DB 307 RPLPPLP 313

RESULT 7

D88088
protein B0454.1 [imported] - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
C/Accession: D88088
R;anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A;Title: Genome sequence of the nematode C. elegans: a platform for investigating biology
A;Reference number: A75000; MUID:99069613; PMID:9851916
A;Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/projects/C.eleg
A;Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A;Accession: D88088
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-348 <STO>
A;Cross-references: UNIPARC:UPI000017A57B; GB:chr_II; PIDN:AB70947.1; PID:G2429396; GSPD
A;Gene: B0454.1
A;Map position: 2

Query Match 85.4%; Score 41; DB 2; Length 348;
Best Local Similarity 87.5%; Pred. No. 69;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 PLPPLPXP 12
DB 37 PLPPLPPP 44

RESULT 8

S03116
 Gene 33 protein, hepatic - rat
 C/Species: Rattus norvegicus (Norway rat)
 C/Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 09-Jul-2004
 C/Accession: S03116; S03402; B30568; S08283
 R/Chrapkiewicz, N.B.; Davis, C.M.; Chu, D.T.W.; Caldwell, C.M.; Granner, D.K.
 Nucleic Acids Res. 17, 6651-6667, 1989
 A/Title: Rat gene 33: analysis of its structure, messenger RNA and basal promoter activity
 A/Reference number: S03116; MUID:89385990; PMID:2780291
 A/Accession: S03116
 A/Molecule type: DNA
 A/Residues: 1-459 <CHR>
 A/Cross-references: UNIPROT:P05432; UNIPARC:UPI000012F0FD; EMBL:X07266; NID:G57568; PIDN:
 R/Lee, K.L.; Makkintje, A.; Ch'ang, L.Y.; Kenney, F.T.
 Arch. Biochem. Biophys. 269, 106-113, 1989
 A/Title: Molecular cloning and analysis of full-length cDNAs cognate to a rat gene under
 A/Reference number: S03402; MUID:89133523; PMID:2916834
 A/Accession: S03402
 A/Molecule type: mRNA
 A/Residues: 1-459 <LEE>
 A/Cross-references: UNIPARC:UPI000012F0FD
 A/Note: the authors translated the codon GGA for residue 18 as Lys, TAC for residues 192
 A/Accession: B30568
 A/Status: preliminary; not compared with conceptual translation
 A/Molecule type: mRNA
 A/Residues: 1-17, 'K', '19-66, 143-191, 'T', 193-301, 'L', 303-310, 'L', 312-395, 'L', 397-409, 'L', 4
 A/Cross-references: UNIPARC:UPI000017C912
 R/Lee, K.L.; Makkintje, A.; Ch'ang, L.Y.; Kenney, F.T.
 Arch. Biochem. Biophys. 276, 554, 1990
 A/Reference number: S08283
 A/Contents: annotation
 A/Note: this is a revision of the nucleotide translation of residues 18, 192, 302, 311,
 C/Genetics:
 A/Gene: 33
 C/Keywords: alternative splicing; liver
 F/1-459/Product: Gene 33 protein, long form #status predicted <MAT1>
 F/1-66,143-459/Product: Gene 33 protein, short form #status predicted <MAT2>

Query Match 85.4%; Score 41; DB 2; Length 459;
 Best Local Similarity 100.0%; Pred. No. 92; Mismatches 0; Indels 0; Gaps 0;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 RPLPPLP 10
 Db 150 RPLPPLP 156

RESULT 9
 T46917
 hypothetical protein DKFZp762K137.1 - human (fragment)
 C/Species: Homo sapiens (man)
 C/Date: 17-Mar-2000 #sequence_revision 17-Mar-2000 #text_change 09-Jul-2004
 R/Octenwälder, B.; Obermaier, B.; Mewes, H.W.; Weil, B.; Wiemann, S.
 submitted to the Protein Sequence Database, February 2000
 A/Reference number: Z24136
 A/Accession: T46917
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-477 <AAA>
 A/Cross-references: UNIPROT:Q9NSQ8; UNIPARC:UPI000006CEC3; EMBL:AL157482
 A/Experimental source: adult melanoma (MeWo cell line); clone DKFZp762K137
 C/Genetics:
 A/Note: DKFZp762K137.1

Query Match 85.4%; Score 41; DB 2; Length 477;
 Best Local Similarity 100.0%; Pred. No. 96; Mismatches 0; Indels 0; Gaps 0;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 RPLPPLP 10
 Db 228 RPLPPLP 234

RESULT 10

A46101
 protein-tyrosine-phosphatase (EC 3.1.3.48) nonreceptor type PTP61F, short splice form -
 C/Species: Drosophila melanogaster
 C/Date: 08-May-1995 #sequence_revision 12-May-1995 #text_change 09-Jul-2004
 R/McLaughlin, S.; Dixon, J.E.
 J. Biol. Chem. 268, 6839-6842, 1993
 A/Title: Alternative splicing gives rise to a nuclear protein tyrosine phosphatase in Dr
 A/Reference number: A46101; MUID:93216607; PMID:8463208
 A/Accession: A46101
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-535 <MCL>
 A/Cross-references: UNIPROT:Q9W0G1; UNIPARC:UPI000016BD13; GB:L11251; NID:G290265; PIDN:
 A/Note: authors translated the codon TTC for residue 382 as Ile, and CGA for residue 528
 C/Genetics:

A/Gene: FlyBase:Ptp61F
 A/Cross-references: FlyBase:FBgn0003138
 C/Superfamily: protein-tyrosine-phosphatase, nonreceptor type PTP61F; protein-tyrosine-ph
 C/Keywords: alternative splicing; phosphoprotein; phosphoric monoester hydrolase; tyrosin
 F/60-285/Domain: protein-tyrosine-phosphatase homology <PTP>
 F/237/Active site: Cys (phosphocysteine intermediate) #status predicted
 F/243/Binding site: substrate phosphate (Arg) #status predicted

Query Match 85.4%; Score 41; DB 2; Length 535;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02; Mismatches 0; Indels 0; Gaps 0;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 RPLPPLP 10
 Db 391 RPLPPLP 397

RESULT 11

B46101
 protein-tyrosine-phosphatase (EC 3.1.3.48), nonreceptor type PTP61F, long splice form -
 C/Species: Drosophila melanogaster
 C/Date: 08-May-1995 #sequence_revision 12-May-1995 #text_change 09-Jul-2004
 R/McLaughlin, S.; Dixon, J.E.
 J. Biol. Chem. 268, 6839-6842, 1993
 A/Title: Alternative splicing gives rise to a nuclear protein tyrosine phosphatase in Dr
 A/Reference number: A46101; MUID:93216607; PMID:8463208
 A/Accession: B46101
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-548 <MCL>
 A/Cross-references: UNIPROT:Q9W0G1; UNIPARC:UPI000016BD12; GB:L11251
 A/Note: authors translated the codon TTC for residue 382 as Ile
 C/Genetics:

A/Gene: FlyBase:Ptp61F
 A/Cross-references: FlyBase:FBgn0003138
 C/Superfamily: protein-tyrosine-phosphatase, nonreceptor type PTP61F; protein-tyrosine-ph
 C/Keywords: alternative splicing; phosphoprotein; phosphoric monoester hydrolase; tyrosin
 F/60-285/Domain: protein-tyrosine-phosphatase homology <PTP>
 F/237/Active site: Cys (phosphocysteine intermediate) #status predicted
 F/243/Binding site: substrate phosphate (Arg) #status predicted

Query Match 85.4%; Score 41; DB 2; Length 548;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02; Mismatches 0; Indels 0; Gaps 0;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 RPLPPLP 10
 Db 391 RPLPPLP 397

RESULT 12

T26064
 hypothetical protein W01G7.1 - Caenorhabditis elegans
 C/Species: Caenorhabditis elegans

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T26064

R;Percy, C.

submitted to the EMBL Data Library, October 1996

A;Reference number: Z20146

A;Accession: T26064

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-627 <WIL>

A;Cross-references: UNIPROT:Q9XVH7; UNIPARC:UPI00000798B; EMBL:Z81135; PIDN: CAB03453.1;

A;Experimental source: clone W01G7

C;Genetics:

A;Gene: CESP:W01G7.1

A;Map position: 2

A;Introns: 82/2; 121/1; 182/2; 588/3

C;Superfamily: Caenorhabditis elegans hypothetical protein W01G7.1

Query Match 85.4%; Score 41; DB 2; Length 627;
Best Local Similarity 87.5%; Pred. No. 1.3e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 PLPLPLXP 12

Db 513 PLPLPLPP 520

RESULT 13

TMVRR

protein-tyrosine kinase (EC 2.7.1.112) fgr - feline sarcoma virus (strain Gardner-Rasheed C;Species: feline sarcoma virus

A;Note: host Felis sp. (cat)

C;Date: 27-Nov-1995 #sequence_revision 26-May-1995 #text_change 09-Jul-2004

R;Accession: A00653; A03937

R;Naharro, G.; Robbins, K.C.; Reddy, E.P.

Science 223, 63-66, 1984

A;Title: Gene product of v-fgr onc: hybrid protein containing a portion of actin and a t

A;Reference number: A00653; MUID: 84097512; PMID: 6318314

A;Accession: A00653

A;Molecule type: DNA

A;Residues: 1-663 <NAH>

A;Cross-references: UNIPROT:P00544; UNIPARC:UPI000017101E; GB:X00255; GB:K01487; NID: g61

A;Note: the authors translated the codon GAT for residue 14 as Glu

C;Comment: This protein is synthesized as a gag-fgr polypeptide.

C;Genetics:

A;Gene: fgr

C;Superfamily: feline sarcoma virus protein-tyrosine kinase fgr; protein kinase homology

C;Keywords: ATP; autophosphorylation; oncogene; phosphoprotein; phosphotransferase; poly

F;1-118/Region: gag polypeptide similarity

F;141-268/Region: actin similarity

F;285-382/Domain: SH2 homology <SH2>

F;402-660/Domain: protein kinase homology <KIN>

F;410-418/Region: protein kinase ATP-binding motif

F;432/Active site: Lys #status predicted

Query Match 85.4%; Score 41; DB 1; Length 663;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 RPLPLPLP 10

Db 131 RPLPLPLP 137

RESULT 14

F84870

hypothetical protein At2g43800 [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004

C;Accession: F84870

R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Unayam, L.; Tallon, L.

euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.

Nature 402, 761-768, 1999

A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID: 20083487; PMID: 10617197

A;Accession: F84870

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-894 <STO>

A;Cross-references: UNIPROT:O22824; UNIPARC:UPI000000A2AE1; GB:AE002093; NID: g2281090; PI:

C;Genetics:

A;Gene: At2g43800

A;Map position: 2

Query Match 85.4%; Score 41; DB 2; Length 894;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 RPLPLPLP 10

Db 289 RPLPLPLP 295

RESULT 15

T16346

hypothetical protein F42C5.10 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004

C;Accession: T16346

R;Du, Z.

submitted to the EMBL Data Library, November 1995

A;Description: The sequence of C. elegans cosmid F42C5.

A;Reference number: Z18497

A;Accession: T16346

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-1225 <DUZ>

A;Cross-references: UNIPROT:Q20330; UNIPARC:UPI000017B9B3; EMBL:U40799; NID: g1065935; PI:

C;Genetics:

A;Gene: CESP:F42C5.10

A;Introns: 11/3; 280/2; 312/3; 646/1; 786/1; 869/1; 1034/1; 1080/2; 1108/1; 1170/1

Query Match 85.4%; Score 41; DB 2; Length 1225;
Best Local Similarity 87.5%; Pred. No. 2.6e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 PLPLPLXP 12

Db 491 PLPLPLPP 498

RESULT 16

T36365

proline-rich protein SCE94.05 - Streptomyces coelicolor

C;Species: Streptomyces coelicolor

C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004

C;Accession: T36365

R;Oliver, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, April 1999

A;Reference number: Z21573

A;Accession: T36365

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-134 <OLI>

A;Cross-references: UNIPROT:Q9X8L8; UNIPARC:UPI000000DAPBA; EMBL:AL049628; PIDN: CAB40854.

A;Experimental source: strain A3(2)

C;Genetics:

A;Gene: SCOEDB:SCE94.05

Query Match 83.3%; Score 40; DB 2; Length 134;
Best Local Similarity 77.8%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RPLPLPLXP 12

Db 102 RPLPLSPPP 110

RESULT 17

GNVOL2

genome-linked protein - potato leaf roll virus
C/Species: potato leaf roll virus
A/Note: host Solanum tuberosum (potato)
C/Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 09-Jul-2004
C/Accession: JQ0002; B60080
R/Kawchuk, L.M.; Martin, R.R.; Rochon, D.M.; McPherson, J.
J. Gen. Virol. 70, 783-788, 1989
A/Title: Identification and characterization of the potato leafroll virus putative coat
A/Reference number: A92803; MUID:89279259; PMID:2732704
A/Accession: JQ0002
A/Molecule type: mRNA
A/Residues: 1-156 <KAW>
A/Cross-references: UNIPROT:PI7523; UNIPARC:UPI00000005FE; GB:D13753; NID:G222290; PIDN:
R/Facke, E.; Sarkar, S.; Salamini, F.; Rohde, W.
Arch. Virol. 105, 153-163, 1989
A/Title: Cloning of the gene for the capsid protein of potato leafroll virus.
A/Reference number: A60080; MUID:89321807; PMID:2751428
A/Accession: B60080
A/Molecule type: mRNA
A/Residues: 1-9, 'G', 11-40, 'V', 42-120, 'A', 122-156 <TAC>
A/Cross-references: UNIPARC:UPI0000002D3ED
C/Species: potato leaf roll virus
C/Superfamily: potato leaf roll virus genome-linked protein
C/Keywords: genome-linked protein

Query Match 83.3%; Score 40; DB 1; Length 156;
Best Local Similarity 77.8%; Pred. No. 41;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 RPLPPLPXP 12
|||||
Db 114 RPLPPPPVP 122

RESULT 18

GNVOL1

genome-linked protein - potato leaf roll virus (strain 1)
C/Species: potato leaf roll virus
A/Note: host Solanum tuberosum (potato)
C/Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 09-Jul-2004
C/Accession: JA0121; S36689
R/Mayo, M.A.; Robinson, D.J.; Jolly, C.A.; Hyman, L.
J. Gen. Virol. 70, 1037-1051, 1989
A/Title: Nucleotide sequence of potato leafroll luteovirus RNA.
A/Reference number: JA0119; MUID:89279282; PMID:2732710
A/Accession: JA0121
A/Molecule type: genomic RNA
A/Residues: 1-156 <MAY>
A/Cross-references: UNIPROT:PI7524; UNIPARC:UPI0000138CCE; EMBL:X14600; NID:G222293; PID
A/Note: the nucleotide sequence was submitted to the EMBL Data Library, March 1989
C/Superfamily: potato leaf roll virus genome-linked protein
C/Keywords: genome-linked protein

Query Match 83.3%; Score 40; DB 1; Length 156;
Best Local Similarity 77.8%; Pred. No. 41;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 RPLPPLPXP 12
|||||
Db 114 RPLPPPPVP 122

RESULT 19

GNVOL4

genome-linked protein - potato leaf roll virus (strain Wageningen)
N/Alternate names: VPg protein
C/Species: potato leaf roll virus
A/Note: host Solanum tuberosum (potato)
C/Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 09-Jul-2004
C/Accession: S03550

R/van der Wilk, F.; Huisman, M.J.; Cornelissen, B.J.C.; Hutterings, H.; Goldbach, R.
FEBS Lett. 245, 51-56, 1989
A/Title: Nucleotide sequence and organization of potato leafroll virus genomic RNA.
A/Reference number: S03546; MUID:89171329; PMID:2466700
A/Accession: S03550
A/Molecule type: genomic RNA
A/Residues: 1-156 <VAN>
A/Cross-references: UNIPROT:PI1625; UNIPARC:UPI0000138CCF; EMBL:Y07496; NID:G61198; PIDN:
C/Superfamily: potato leaf roll virus genome-linked protein
C/Keywords: genome-linked protein

Query Match 83.3%; Score 40; DB 1; Length 156;
Best Local Similarity 77.8%; Pred. No. 41;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 RPLPPLPXP 12
|||||
Db 114 RPLPPPPVP 122

RESULT 20

T15811

hypothetical protein C46H3.1 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C/Accession: T15811
R/Ravello, T.

submitted to the EMBL Data Library, November 1995
A/Description: The sequence of C. elegans cosmid C46H3.
A/Reference number: Z18409
A/Accession: T15811
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-215 <FAV>
A/Cross-references: UNIPROT:Q18661; UNIPARC:UPI000007AB2C; EMBL:U41271; NID:G1086855; PTI
C/Genetics:
A/Gene: CESP:C46H3.1
A/Introns: 19/1; 71/1; 99/1; 126/1

Query Match 83.3%; Score 40; DB 2; Length 215;
Best Local Similarity 75.0%; Pred. No. 57;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 PLPPLPXP 12
|||||
Db 104 PLPPMPVP 111

RESULT 21

T16669

hypothetical protein R03H10.2 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 04-Mar-2000
C/Accession: T16669
R/Wilcox, L.

submitted to the EMBL Data Library, June 1995
A/Description: The sequence of C. elegans cosmid R03H10.
A/Reference number: Z18557
A/Accession: T16669
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-224 <WIL>
A/Cross-references: UNIPARC:UPI000017941B; EMBL:U29382; NID:G868226; PID:G868235; PIDN:A
A/Experimental source: strain Bristol N2
C/Genetics:
A/Gene: CESP:R03H10.2

A/Introns: 47/2; 68/3; 127/2; 159/2; 179/2; 209/2
C/Superfamily: Caenorhabditis elegans hypothetical protein R03H10.2

Query Match 83.3%; Score 40; DB 2; Length 224;
Best Local Similarity 75.0%; Pred. No. 60;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 PLPPLXP 12
|||||
Db 164 PLPPVPSP 171

RESULT 22

T29543
hypothetical protein F48C1.6 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T29543
R;Gattung, S.; Le, T.T.
submitted to the EMBL Data Library, April 1997
A;Description: The sequence of C. elegans cosmid F48C1.
A;Reference number: 220638
A;Accession: T29543
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-245 <GAT>
A;Cross-references: UNIPROT:O01572; UNIPARC:UPI000007781C; EMBL:U97015; PIDN:AAB52342.1.
A;Experimental source: strain Bristol N2; clone F48C1
C;Genetics:
A;Gene: CESP:F48C1.6
A;Map position: 1
A;Introns: 52/3; 96/2; 123/3; 161/3
C;Superfamily: Caenorhabditis elegans hypothetical protein F48C1.6

Query Match 83.3%; Score 40; DB 2; Length 245;
Best Local Similarity 66.7%; Pred. No. 66;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 RPLPLXP 12
|||||
Db 208 RPSPPSP 216

RESULT 23

C70551
hypothetical protein RV2507 - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C;Accession: C70551
R;Cole, S.T.; Broech, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: C70551
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-273 <COL>
A;Cross-references: UNIPROT:O06170; UNIPARC:UPI00001652FA; GB:Z95556; GB:AL123456; NID:9
A;Experimental source: strain H37RV
C;Genetics:
A;Gene: RV2507

Query Match 83.3%; Score 40; DB 2; Length 273;
Best Local Similarity 75.0%; Pred. No. 73;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 PLPPLXP 12
|||||
Db 144 PLPPMPGP 151

RESULT 24

T41395
probable dna polymerase alpha-primase associated subunit - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Dec-2004
C;Accession: T41395

R;Wood, V.; Rajandream, M.A.; Barrell, B.G.; Murphy, L.; Harris, D.
submitted to the EMBL Data Library, May 1998
A;Reference number: Z21991
A;Accession: T41395
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-574 <WOO>
A;Cross-references: UNIPROT:O74946; UNIPARC:UPI000012977E; EMBL:AL023704; PIDN:CAA19261.
A;Experimental source: strain 972h-; cosmid C553
C;Genetics:
A;Gene: SPDB:SPCC553.09C
A;Map position: 3
A;Introns: 89/2; 415/3; 518/3
C;Superfamily: DNA polymerase alpha, subunit B

Query Match 83.3%; Score 40; DB 2; Length 574;
Best Local Similarity 75.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 PLPPLXP 12
|||||
Db 278 PLPPPLPP 285

RESULT 25

C81594
hypothetical protein CP0282 [imported] - Chlamydomophila pneumoniae (strain AR39)
C;Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C;Accession: C81594
R;Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey, I.; C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg, E.; C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg, E.
Nucleic Acids Res. 28, 1397-1406, 2000
A;Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A;Reference number: A81500; MUID:20150255; PMID:10684935
A;Accession: C81594
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-775 <REA>
A;Cross-references: UNIPROT:Q9JUS20; UNIPARC:UPI000000D2F9E; GB:AE002189; GB:AE002161; NID:
A;Experimental source: strain AR39, HL cells
C;Genetics:
A;Gene: CP0282

Query Match 83.3%; Score 40; DB 2; Length 775;
Best Local Similarity 75.0%; Pred. No. 2.2e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 PLPPLXP 12
|||||
Db 599 PLPPVPSP 606

RESULT 26

D86549
hypothetical protein CPj0472 [imported] - Chlamydomophila pneumoniae (strain J138)
C;Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C;Accession: D86549
R;Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Is
Nucleic Acids Res. 28, 2311-2314, 2000
A;Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.
A;Reference number: A86491; MUID:20330349; PMID:108711362
A;Accession: D86549
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-775 <STO>
A;Cross-references: UNIPROT:Q9JUS20; UNIPARC:UPI000000D2F9E; GB:BA000008; NID:98978842; PI:
A;Experimental source: strain J138
C;Genetics:
A;Gene: CPj0472

Query Match 83.3%; Score 40; DB 2; Length 775;

Best Local Similarity 75.0%; Pred. No. 2.2e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 PLPPLXP 12
|||||
Db 599 PLPPVPSP 606

RESULT 27
B72074
hypothetical protein - Chlamydomophila pneumoniae (strain CWL029)
C/Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae
C/Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C/Accession: B72074
R/Kalman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.;
Nature Genet. 21, 385-389, 1999
A/Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.
A/Reference number: A72000; MUID:9206606; PMID:10192388
A/Accession: B72074
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-775 <ARN>
A/Cross-references: UNIPROT:Q92879; UNIPARC:UPI00000470F0; GB:AE001632; GB:AE001363; NID
A/Experimental source: strain CWL029
C/Genetics:
A/Gene: CPn0472

Query Match 83.3%; Score 40; DB 2; Length 775;
Best Local Similarity 75.0%; Pred. No. 2.2e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 PLPPLXP 12
|||||
Db 599 PLPPVPSP 606

RESULT 28
S60735
splicing factor SF3a 120K chain - human
C/Species: Homo sapiens (man)
C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C/Accession: S60735; S60733
R/Kraemer, A.; Mulhauser, F.; Wersig, C.; Groening, K.; Bilbe, G.
RNA 1, 260-272, 1995
A/Title: Mammalian splicing factor SF3a120 represents a new member of the SURP family of
A/Reference number: S60733; MUID:56079958; PMID:7489498
A/Accession: S60735
A/Status: nucleic acid sequence not shown
A/Molecule type: mRNA
A/Residues: 1-793 <KRA>
A/Cross-references: UNIPROT:Q15459; UNIPARC:UPI0000000C88; EMBL:X85237; NID:g899297; PID
A/Accession: S60733
A/Molecule type: protein
A/Residues: 51-62;82-94;270-275;397-414;448-463 <KRA2>
A/Cross-references: UNIPARC:UPI0000174C3E; UNIPARC:UPI00000174C3F; UNIPARC:UPI00000174C40;
C/Genetics:
A/Gene: GDB:SF3A120; PRP21; SAP114
A/Cross-references: GDB:9955873
A/Map position: 22q12.1-22qter
C/Superfamily: human splicing factor SF3a 120K chain; ubiquitin homology
C/Keywords: pre-mRNA splicing
F/714-790/Domain: ubiquitin homology <UBH>

Query Match 83.3%; Score 40; DB 1; Length 793;
Best Local Similarity 66.7%; Pred. No. 2.2e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 RPLPPLXP 12
|||||
Db 399 KPLPPAPAP 407

RESULT 29

G30010
hypothetical ORF-8 protein - Leishmania tarentolae mitochondrion
C/Species: mitochondrion Leishmania tarentolae
C/Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 07-Dec-1999
C/Accession: G30010
R/de la Cruz, V.F.; Neckelmann, N.; Simpson, L.
J. Biol. Chem. 259, 15136-15147, 1984
A/Title: Sequences of six genes and several open reading frames in the kinetoplast maxic
A/Reference number: A22848; MUID:85079995; PMID:6096360
A/Accession: G30010
A/Molecule type: DNA
A/Residues: 1-93
A/Cross-references: UNIPARC:UPI000017B58D; GB:M10126
C/Genetics:
A/Genome: mitochondrion
A/Genetic code: SGC6
C/Keywords: mitochondrion

Query Match 81.2%; Score 39; DB 2; Length 93;
Best Local Similarity 75.0%; Pred. No. 33;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 PLPPLXP 12
|||||
Db 43 PLPPFPSP 50

RESULT 30
F83532
conserved hypothetical protein PA0915 [imported] - Pseudomonas aeruginosa (strain PA01)
C/Species: Pseudomonas aeruginosa
C/Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C/Accession: F83532
R/Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Bz:
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim:
; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A/Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A/Reference number: A82950; MUID:20437337; PMID:10984043
A/Accession: F83532
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-153 <STO>
A/Cross-references: UNIPROT:Q91542; UNIPARC:UPI000000C51B6; GB:AE004525; GB:AE004091; NID
A/Experimental source: strain PA01
C/Genetics:
A/Gene: PA0915

Query Match 81.2%; Score 39; DB 2; Length 153;
Best Local Similarity 77.8%; Pred. No. 55;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 RPLPPLXP 12
|||||
Db 75 RPLPPLP 83

RESULT 31
T19341
hypothetical protein Cl6D6.1 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C/Accession: T19341
R/Gardner, A.
submitted to the EMBL Data Library, November 1996
A/Reference number: Z19111
A/Accession: T19341
A/Status: preliminary; translated from GB/EMBL/DDBJ
A/Molecule type: DNA
A/Residues: 1-173 <WIL>
A/Cross-references: UNIPROT:O62063; UNIPARC:UPI000007A889; EMBL:Z81472; PIDN:CAB03889.1;
C/Experimental source: clone Cl6D6
C/Genetics:

A;Gene: CESP:C16D6.1
A;Map position: X
A;Introns: 42/1

Query Match 81.2%; Score 39; DB 2; Length 173;
Best Local Similarity 85.7%; Pred. No. 63;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 4 RPLPPLP 10
|||:||||
Db 121 RMPPLP 127

RESULT 32

A70856

probable lppZ protein - Mycobacterium tuberculosis (strain H37RV)

C;Species: Mycobacterium tuberculosis

C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004

C;Accession: A70856

R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998

A;Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A;Reference number: A70500; MUID:98295987; PMID:9634230

A;Accession: A70856

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-373 <COL>

A;Cross-references: UNIPROT:O53253; UNIPARC:UPI00000318FB; GB:AL021287; GB:AL123456; NID

A;Experimental source: strain H37RV

C;Genetics:

A;Gene: lppZ

Query Match 81.2%; Score 39; DB 2; Length 373;
Best Local Similarity 75.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 5 PLPLPXP 12
|||:|
Db 54 PLPPVPP 61

RESULT 33

S60975

hypothetical protein YNL152w - yeast (Saccharomyces cerevisiae)

N;Alternate names: hypothetical protein N1765

C;Species: Saccharomyces cerevisiae

C;Date: 15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 09-Jul-2004

C;Accession: S60975; S63104; S63822

R;Naar, F.; Becam, A.M.; Herbert, C.J.

submitted to the EMBL Data Library, October 1995

A;Description: The sequence of 36.8 kb from the left arm of chromosome XIV reveals 24 complete tunic dystrophy kinase.

A;Reference number: S60958

A;Accession: S60975

A;Molecule type: DNA

A;Residues: 1-409 <NAS>

A;Cross-references: UNIPROT:P53901; UNIPARC:UPI000013BB88; EMBL:X92517; NID:g1050783; PI

R;Naar, F.; Becam, A.M.; Herbert, C.

submitted to the Protein Sequence Database, April 1996

A;Reference number: S62967

A;Accession: S63104

A;Molecule type: DNA

A;Residues: 1-409 <NAS>

A;Cross-references: UNIPARC:UPI000013BB88; EMBL:Z71428; NID:g1302109; PID:e2339813; PID:9

A;Experimental source: strain S288C

R;Naar, F.; Becam, A.M.; Herbert, C.J.

Yeast 12, 169-175, 1996

A;Title: The sequence of 36.8 kb from the left arm of chromosome XIV reveals 24 complete dystrophy kinase.

A;Reference number: S63805; MUID:96287653; PMID:8686380

A;Accession: S63822

A;Status: nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-409 <NAP>

A;Cross-references: UNIPARC:UPI000013BB88; EMBL:X92517; NID:g1050783; PIDN:CAA63287.1; P

A;Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1995

C;Genetics:

A;Cross-references: SGD:S0005096

A;Map position: 14L

A;Note: YNL152w

Query Match 81.2%; Score 39; DB 2; Length 409;
Best Local Similarity 85.7%; Pred. No. 1.5e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 4 RPLPPLP 10
|||:||||
Db 159 RPIPLP 165

RESULT 34

A34794

B-cell CLL/lymphoma 3 (BCL3) protein - human

C;Species: Homo sapiens (man)

C;Date: 13-Jul-1990 #sequence_revision 13-Jul-1990 #text_change 31-Dec-2004

C;Accession: A34794

R;Ohno, H.; Takimoto, G.; McKeithan, T.W.

Cell 60, 991-997, 1990

A;Title: The candidate proto-oncogene bcl-3 is related to genes implicated in cell linea

A;Reference number: A34794; MUID:90199880; PMID:2180580

A;Accession: A34794

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-446 <OHN>

A;Cross-references: UNIPROT:P20749; UNIPARC:UPI000012682B; GB:M31732; NID:g179375; PIDN:

C;Genetics:

A;Gene: GDB:BCL3; BCL4: D19S37

A;Cross-references: GDB:I20561; OMIM:109560

A;Map position: 19q13.2-19q13.2

C;Superfamily: EGF homology

C;Keywords: transcription regulation

Query Match 81.2%; Score 39; DB 2; Length 446;
Best Local Similarity 66.7%; Pred. No. 1.7e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RPLPPLXP 12
|||:|
Db 435 RPVPSPAP 443

RESULT 35

A75464

probable zinc metall proteinase - Deinococcus radiodurans (strain R1)

C;Species: Deinococcus radiodurans

C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004

C;Accession: A75464

R;White, O.; Eissen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;

M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma

S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.

Science 286, 1571-1577, 1999

A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.

A;Reference number: A75250; MUID:20036896; PMID:10567266

A;Accession: A75464

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-472 <WHI>

A;Cross-references: UNIPROT:Q9RVZ5; UNIPARC:UPI00000D3D6E; GB:AE001941; GB:AE000513; NID

A;Experimental source: strain R1

C;Genetics:

A;Gene: DR0875

A;Map position: 1

Query Match 81.2%; Score 39; DB 2; Length 472;
Best Local Similarity 77.8%; Pred. No. 1.8e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 RPLPPLXP 12
| | | | | |
Db 459 RTLPPLPEP 467

RESULT 36
C49507
potassium channel Kv1.5, form 3 - mouse
C/Species: Mus musculus (house mouse)
C/Date: 10-Nov-1995 #sequence_revision 10-Nov-1995 #text_change 17-Nov-2000
C/Accession: C49507
R/Attali, B.; Lesage, F.; Zilliani, P.; Guillemare, E.; Honore, E.; Waldmann, R.; Hugnot, J. Biol. Chem. 268, 24283-24289, 1993
A/Title: Multiple mRNA isoforms encoding the mouse cardiac Kv1-5 delayed rectifier K(+)
A/Reference number: A49507; MUID:94043264; PMID:8226976
A/Accession: C49507
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-514 <ATT>
A/Cross-references: UNIPARC:UPI00001779F5; GB:L22218
C/Superfamily: potassium channel protein drk1
C/Keywords: alternative splicing

Query Match 81.2%; Score 39; DB 2; Length 514;
Best Local Similarity 85.7%; Pred. No. 2e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 RPLPPLP 10
| | | | | |
Db 65 RPLPPMP 71

RESULT 37
A49507
potassium channel Kv1.5 - mouse
C/Species: Mus musculus (house mouse)
C/Date: 10-Nov-1995 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
C/Accession: A49507; B49507
R/Attali, B.; Lesage, F.; Zilliani, P.; Guillemare, E.; Honore, E.; Waldmann, R.; Hugnot, J. Biol. Chem. 268, 24283-24289, 1993
A/Title: Multiple mRNA isoforms encoding the mouse cardiac Kv1-5 delayed rectifier K(+)
A/Reference number: A49507; MUID:94043264; PMID:8226976
A/Accession: A49507
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-602 <ATT>
A/Cross-references: UNIPROT:Q61762; UNIPARC:UPI0000028EB9; GB:L22218; NID:G435603; PIDN:
A/Accession: B49507
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 201-602 <AT2>
A/Cross-references: UNIPARC:UPI000002A65A; GB:L22218
C/Superfamily: potassium channel protein drk1
C/Keywords: alternative splicing

Query Match 81.2%; Score 39; DB 2; Length 602;
Best Local Similarity 85.7%; Pred. No. 2.3e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 RPLPPLP 10
| | | | | |
Db 65 RPLPPMP 71

RESULT 38
T49494
condensin complex component cnd2 related protein [imported] - Neurospora crassa
N/Alternate names: protein B14D6.460
C/Species: Neurospora crassa

C/Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 09-May-2004
C/Accession: T49494
R/Schulte, U.; Algm, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura, submitted to the Protein Sequence Database, May 2000
A/Reference number: Z25022
A/Accession: T49494
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-832 <SCH>
A/Cross-references: UNIPARC:UPI00001798B1; EMBL:AL356173; GSPDB:GN00116; NCSP:B14D6.460
A/Experimental source: BAC clone B14D6; strain OR74A
C/Genetics:
A/Gene: NCSP:B14D6.460
A/Map position: 6
A/Introns: 122/3
C/Superfamily: chromosome condensation complex condensin, subunit H

Query Match 81.2%; Score 39; DB 2; Length 832;
Best Local Similarity 77.8%; Pred. No. 3.2e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 RPLPPLXP 12
| | | | | |
Db 754 RRLPPLPTP 762

RESULT 39
T31742
hypothetical protein C05C8.4 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C/Accession: T31742
R/Sammons, L.; Wohldmann, P.
submitted to the EMBL Data Library, July 1997
A/Description: The sequence of C. elegans cosmid C05C8.
A/Reference number: Z21078
A/Accession: T31742
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-1106 <SAM>
A/Cross-references: UNIPROT:O16310; UNIPARC:UPI00000749A5; EMBL:AF016430; PIDN:AAB65371.1
A/Experimental source: strain Bristol N2; clone C05C8
C/Genetics:
A/Gene: C05C8.4
A/Map position: 5
A/Introns: 25/3; 78/3; 117/1; 245/1; 591/1; 787/1; 1008/2

Query Match 81.2%; Score 39; DB 2; Length 1106;
Best Local Similarity 66.7%; Pred. No. 4.4e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 RPLPPLXP 12
| | | | | |
Db 1038 RMPPPMP 1046

RESULT 40
T42230
AF4 protein - mouse (fragment)
N/Alternate names: serine/proline-rich FEL protein
C/Species: Mus musculus (house mouse)
C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C/Accession: T42230
R/Baskaran, K.; Erfurth, F.; Taborn, G.; Copeland, N.G.; Gilbert, J.; Jenkins, N.A.; Iann submitted to the EMBL Data Library, July 1997
A/Description: Cloning and developmental expression of the murine homolog of the acute le A/Reference number: Z22090
A/Accession: T42230
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-1211 <BAS>
A/Cross-references: UNIPROT:O35233; UNIPARC:UPI0000029466; EMBL:AF013131; NID:G2582018; I A/Experimental source: strain CD1

C;Genetics:
A;Gene: Af4
A;Map position: 5

Query Match 81.2%; Score 39; DB 2; Length 1211;
Best Local Similarity 66.7%; Pred. No. 4.8e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 RPLPPLXP 12
||| ||| |||
Db 655 RPKPPVPTP 663

RESULT 41

T42625
AF-4 protein - mouse
C;Species: Mus musculus (house mouse)
C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jul-2004
C;Accession: T42625
R;Isnard, P.; Depetris, D.; Mattei, M.G.; Ferrier, P.; Djabali, M.
Mamm. Genome 9, 1065-1068, 1998
A;Title: CDNA cloning, expression and chromosomal localization of the murine AF-4 gene
A;Reference number: 222176; MUID:99099257; PMID:9880680
A;Accession: T42625
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-1217 <ISN>
A;Cross-references: UNIPROT:O88573; UNIPARC:UPI0000029921; EMBL:AF074266; NID:g3328189;
A;Experimental source: strain M; thymus

Query Match 81.2%; Score 39; DB 2; Length 1217;
Best Local Similarity 66.7%; Pred. No. 4.8e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 RPLPPLXP 12
||| ||| |||
Db 660 RPKPPVPTP 668

RESULT 42

B45344
probable nuclear antigen - suid herpesvirus 1 (strain Kaplan)
C;Species: suid herpesvirus 1
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
C;Accession: B45344
R;Vlcek, C.; Kozmik, Z.; Paces, V.; Schirm, S.; Schwyzler, M.
Virology 179, 365-377, 1990
A;Title: Pseudorabies virus immediate-early gene overlaps with an oppositely oriented of
A;Reference number: A45344; MUID:91021039; PMID:2171211
A;Accession: B45344
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-1733 <VLC>
A;Cross-references: UNIPROT:P33485; UNIPARC:UPI00000138ADE; GB:M34651; NID:g334070; PIDN:
C;Superfamily: pseudorabies virus 1 nuclear antigen

Query Match 81.2%; Score 39; DB 1; Length 1733;
Best Local Similarity 77.8%; Pred. No. 7e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 RPLPPLXP 12
||| ||| |||
Db 279 RPPPPPLPPP 287

RESULT 43

B40505
hypothetical protein - suid herpesvirus 1 (strain Indiana-Funkhuser or Becker)
C;Species: suid herpesvirus 1
C;Date: 10-Apr-1992 #sequence_revision 10-Apr-1992 #text_change 09-Jul-2004
C;Accession: B40505
R;Cheung, A.K.
J. Virol. 65, 5260-5271, 1991

A;Title: Cloning of the latency gene and the early protein 0 gene of pseudorabies virus.
A;Reference number: A40505; MUID:91374576; PMID:1654441
A;Accession: B40505
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1958 <CHE>
A;Cross-references: UNIPROT:Q69340; UNIPARC:UPI00000F2308; GB:M57505; NID:g334066; PIDN:
C;Superfamily: pseudorabies virus 1 nuclear antigen

Query Match 81.2%; Score 39; DB 2; Length 1958;
Best Local Similarity 77.8%; Pred. No. 7.9e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 RPLPPLXP 12
||| ||| |||
Db 488 RPPPPPLPPP 496

RESULT 44

S72349
nonstructural polyprotein - eastern equine encephalomyelitis virus
N;Alternate names: nonstructural protein NSP1; nonstructural protein NSP2; nonstructural
C;Species: eastern equine encephalomyelitis virus
C;Date: 04-May-1998 #sequence_revision 15-May-1998 #text_change 09-Jul-2004
C;Accession: S72349
R;Weaver, S.C.; Hagenbaugh, A.; Bellew, L.A.; Netesov, S.V.; Volchkov, V.E.; Chang, G.J.
Virology 197, 375-390, 1993
A;Title: A comparison of the nucleotide sequences of eastern and western equine encephal
A;Reference number: S72349; MUID:94025587; PMID:8105605
A;Accession: S72349
A;Status: preliminary
A;Molecule type: genomic RNA
A;Residues: 1-2493 <WEA>
A;Cross-references: UNIPROT:Q88789; UNIPARC:UPI00000F4983; EMBL:U01034; NID:g393006; PIDN:
A;Note: readthrough of the terminator UGA occurs between the codons AAU for residue 1879
A;Note: the readthrough stopcodon UGA for residue 1879 is translated as X
C;Superfamily: Semliki Forest virus nonstructural protein

Query Match 81.2%; Score 39; DB 2; Length 2493;
Best Local Similarity 66.7%; Pred. No. 1e+03;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 RPLPPLXP 12
||| ||| |||
Db 1670 RPAPPVPVP 1678

RESULT 45

S26372
nonstructural polyprotein - eastern equine encephalomyelitis virus
N;Contains: nonstructural protein NSP1; nonstructural protein NSP2; nonstructural protei
C;Species: eastern equine encephalomyelitis virus
C;Date: 06-Jan-1994 #sequence_revision 24-Jul-1997 #text_change 31-Dec-2004
C;Accession: S26369; S26372
R;Volchkov, V.E.; Volchkov, V.A.; Netesov, S.V.
Mol. Gen. Microbiol. Virusol. 5, 8-15, 1991
A;Title: Complete nucleotide sequence of the eastern equine encephalomyelitis virus geno
A;Reference number: S26369; MUID:91375524; PMID:1896061
A;Accession: S26369
A;Molecule type: genomic RNA
A;Residues: 1-2493 <VOL>
A;Cross-references: UNIPROT:Q66581; UNIPARC:UPI0000178569; EMBL:X63135
A;Note: sequence could not be checked because of bad print in paper
A;Note: readthrough of the terminator UGA occurs between the codons AAU for residue 1878
A;Note: the readthrough stopcodon UGA for residue 1879 is translated as X
A;Accession: S26372
A;Molecule type: mRNA
A;Residues: 1886-2493 <VO2>
A;Cross-references: UNIPARC:UPI00000F0DC1; EMBL:X63135; NID:g59185; PIDN:CAA44847.1; PID
A;Note: sequence could not be checked because of bad print in paper
A;Note: this reading frame extends between two stop codons and does not begin with a sta
C;Keywords: nonstructural protein; polyprotein
F;1-532/Product: nonstructural protein NSP1 #status predicted <NS1>

Query Match 81.2%; Score 39; DB 1; Length 1733;
Best Local Similarity 77.8%; Pred. No. 7e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 RPLPPLXP 12
||| ||| |||
Db 279 RPPPPPLPPP 287

RESULT 43

B40505
hypothetical protein - suid herpesvirus 1 (strain Indiana-Funkhuser or Becker)
C;Species: suid herpesvirus 1
C;Date: 10-Apr-1992 #sequence_revision 10-Apr-1992 #text_change 09-Jul-2004
C;Accession: B40505
R;Cheung, A.K.
J. Virol. 65, 5260-5271, 1991

F/533-1326/Product: nonstructural protein NSP2 #status predicted <NS2>
 F/1327-1878/Product: nonstructural protein NSP3 #status predicted <NS3>
 F/1886-2493/Product: nonstructural protein NSP4 #status predicted <NS4>

Query Match 81.2%; Score 39; DB 2; Length 2493;
 Best Local Similarity 66.7%; Pred. No. 1e+03;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 RPLPPLXP 12
 |||||
 Db 1670 RPAPPVVP 1678

RESULT 46
 E87601
 OmpA family protein [imported] - Caulobacter crescentus
 C/Species: Caulobacter crescentus
 C/Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
 C/Accession: E87601
 R/Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
 B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
 n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
 Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
 A/Title: Complete Genome Sequence of Caulobacter crescentus.
 A/Reference number: A87249; MUID:21173698; PMID:11259647
 A/Accession: E87601
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-191 <STO>
 A/Cross-references: UNIPROT:Q9A4I6; UNIPARC:UPI00000C7879; GB:AE005673; NID:g13424457; E
 C/Genetics:
 A/Gene: CC2845

Query Match 79.2%; Score 38; DB 2; Length 191;
 Best Local Similarity 66.7%; Pred. No. 96;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 RPLPPLXP 12
 |||||
 Db 62 QLPAPPAPLP 70

RESULT 47
 T04353
 DNA binding protein - rice
 C/Species: Oryza sativa (rice)
 C/Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
 C/Accession: T04353
 R/Haing, Y.C.; Teao, C.V.; Chow, T.; Hsieh, J.; Chen, Z.
 submitted to the EMBL Data Library, April 1995
 A/Description: Rice early embryogenesis gene.
 A/Reference number: Z14889
 A/Accession: T04353
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: mRNA
 A/Residues: 1-217 <HSI>
 A/Cross-references: UNIPROT:P93426; UNIPARC:UPI000009C6C6; EMBL:U25283; NID:g1753084; PT
 A/Experimental source: cv. Tainung 67, seed

Query Match 79.2%; Score 38; DB 2; Length 217;
 Best Local Similarity 66.7%; Pred. No. 1.1e+02;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 RPLPPLXP 12
 |||||
 Db 13 RPSPLPAP 21

RESULT 48
 C75539
 conserved hypothetical protein - Deinococcus radiodurans (strain R1)
 C/Species: Deinococcus radiodurans
 C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004

C/Accession: C75539
 R/White, O.; Eisen, J.A.; Heidelberg, J.P.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
 M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
 Science 286, 1571-1577, 1999
 A/Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
 A/Reference number: A75250; MUID:20036896; PMID:10567266
 A/Accession: C75539
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-222 <WHI>
 A/Cross-references: UNIPROT:Q9RXN1; UNIPARC:UPI00000C171C; GB:AE001889; GB:AE000513; NID
 A/Experimental source: strain R1
 C/Genetics:
 A/Gene: DR0279
 A/Map position: 1

Query Match 79.2%; Score 38; DB 2; Length 222;
 Best Local Similarity 85.7%; Pred. No. 1.1e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 RPLPPLP 10
 |||||
 Db 201 RPVPPLP 207

RESULT 49
 A38582
 pollen allergen p1b precursor - perennial ryegrass
 N/Alternate names: 30K allergen
 C/Species: Lolium perenne (perennial ryegrass)
 C/Date: 14-Feb-1992 #sequence_revision 14-Feb-1992 #text_change 07-May-1999
 C/Accession: A38582; S38290
 R/Singh, M.B.; Hough, T.; Theerakulpisut, P.; Avjoglou, A.; Davies, S.; Smith, P.M.; Tayl
 Proc. Natl. Acad. Sci. U.S.A. 88, 1384-1388, 1991
 A/Title: Isolation of cDNA encoding a newly identified major allergenic protein of rye-g
 A/Reference number: A38582; MUID:91142177; PMID:1671715
 A/Accession: A38582
 A/Status: preliminary; not compared with conceptual translation
 A/Molecule type: mRNA
 A/Residues: 1-308 <SIN>
 A/Cross-references: UNIPARC:UPI0000177EPF; GB:M59163
 R/Petersen, A.; Schramm, G.; Becker, W.M.; Schlaak, M.
 Biol. Chem. Hoppe-Seyler 374, 855-861, 1993
 A/Title: Comparison of four grass pollen species concerning their allergens of grass gro
 A/Reference number: S38288; MUID:94092339; PMID:7505588
 A/Accession: S38290
 A/Molecule type: protein
 A/Residues: 26-45 <PET>
 A/Cross-references: UNIPARC:UPI00001474F3
 C/Superfamily: grass pollen allergen IX
 C/Keywords: pollen
 F/1-25/Domain: signal sequence #status predicted <SIG>
 F/26-308/Product: pollen allergen p1b #status experimental <MAT>

Query Match 79.2%; Score 38; DB 2; Length 308;
 Best Local Similarity 66.7%; Pred. No. 1.6e+02;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 RPLPPLXP 12
 |||||
 Db 70 QPWPPLPTP 78

RESULT 50
 A31978
 ADP, ATP carrier protein AAC2 precursor - yeast (Saccharomyces cerevisiae)
 N/Alternate names: adenine nucleotide carrier; protein YBL0421
 C/Species: Saccharomyces cerevisiae
 C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
 C/Accession: A31978; S36419; S45764; S50304; S42759
 R/Lawson, J.E.; Douglas, M.G.
 J. Biol. Chem. 263, 14812-14818, 1988

A;Title: Separate genes encode functionally equivalent ADP/ATP carrier proteins in Sacch
A;Reference number: A31978; MUID:89008354; PMID:2844786
A;Accession: A31978
A;Molecule type: DNA
A;Residues: 1-318 <LAW>
A;Cross-references: UNIPROT:P19239; UNIPARC:UPI000012561E; EMBL:J04021; NID:gi70957; PID
R;Trezequet, V.; le Saux, A.; Lauquin, J.M.
submitted to the EMBL Data Library, August 1993
A;Reference number: S36419
A;Accession: S36419
A;Molecule type: DNA
A;Residues: 1-318 <TRE>
A;Cross-references: UNIPARC:UPI000012561E; EMBL:X74427; NID:G396554; PIDN:CAA52446.1; PI
R;Goffeau, A.; Jonniaux, J.L.; Purnelle, B.; Skala, J.; de Wergifosse, P.; van Dyck, L.
submitted to the Protein Sequence Database, August 1994
A;Reference number: S45745
A;Accession: S45745
A;Molecule type: DNA
A;Residues: 1-318 <GOF>
A;Cross-references: UNIPARC:UPI000012561E; EMBL:Z35791; NID:G536034; PIDN:CAA84850.1; PI
R;van Dyck, L.; Jonniaux, J.L.; de Melo Barreiros, T.; Kleine, K.; Goffeau, A.
Yeast 10, 1663-1673, 1994
A;Title: Analysis of a 17.4 kb DNA segment of yeast chromosome II encompassing the ribos
human proliferation-associated p120 antigen.
A;Reference number: S50299; MUID:95242843; PMID:7725803
A;Accession: S50304
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-318 <VAN>
A;Cross-references: UNIPARC:UPI000012561E; EMBL:X77291; NID:G602888; PIDN:CAA54501.1; PI
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1994
R;Kolarov, J.; Kolarova, N.; Nelson, N.
J. Biol. Chem. 265, 12711-12716, 1990
A;Title: A third ADP/ATP translocator gene in yeast.
A;Reference number: A36592; MUID:90324269; PMID:2165073
A;Accession: S42759
A;Molecule type: DNA
A;Residues: 1-57, 'I', '59-64', 'K', '66-67', 'S', '69-70', 'V', '72-78', 'K', '80-82', 'L', '84-112', 'L', '114-12
A;Cross-references: UNIPARC:UPI0000169AD7; EMBL:M34075; NID:gi70961; PIDN:AAA97484.1; PI
C;Genetics:
A;Gene: SGD:PET9; AAC2; MIPS:YBL030c
A;Cross-references: MIPS:YBL030c; SGD:S0000126
A;Map position: 2L
A;Genome: nuclear
C;Superfamily: ADP, ATP carrier protein; ADP, ATP carrier protein repeat homology
C;Keywords: duplication; homodimer; mitochondrial inner membrane; mitochondrion; transme
F;1-20/Domain: transit peptide (mitochondrion) #status predicted <TNP>
F;20-115/Domain: ADP, ATP carrier protein repeat homology <ACPI>
F;21-318/Product: ADP, ATP carrier protein AAC2 #status predicted <MAT>
F;21-37/Domain: transmembrane #status predicted <TM1>
F;124-218/Domain: ADP, ATP carrier protein repeat homology <ACP2>
F;131-147/Domain: transmembrane #status predicted <TM2>
F;186-202/Domain: transmembrane #status predicted <TM3>
F;217-233/Domain: transmembrane #status predicted <TM4>
F;224-312/Domain: ADP, ATP carrier protein repeat homology <ACP3>

Query Match 79.2%; Score 38; DB 1; Length 318;
Best Local Similarity 75.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 5 PLPPLXP 12
| | | | |
Db 10 PLPPAPAP 17

Search completed: April 6, 2006, 09:39:41
Job time : 48.7895 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 6, 2006, 09:23:24 ; Search time 145.895 Seconds
(without alignments)
58.030 Million cell updates/sec

Title: US-10-632-388-307

Perfect score: 48

Sequence: 1 XXXRPLEPLXP 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Database :

1: uniprot_05.80.*

2: uniprot_sprot.*

3: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	48	100.0	1112	Q41257_GIBZE	Q41257 gibberella
2	47	97.9	613	KCNAS_HUMAN	P22460 homo sapien
3	47	97.9	613	Q4VAU1_HUMAN	Q4VAU1 homo sapien
4	46	95.8	282	1 TNFL6_PIG	Q9BEA8 sus scrofa
5	46	95.8	1794	Q552A2_CRYNE	Q552A2 cryptococcus
6	46	95.8	1811	Q5KN17_CRYNE	Q5KN17 cryptococcus
7	45	93.8	85	1 FDSCP_HUMAN	Q8NFU4 homo sapien
8	45	93.8	85	Q540F3_HUMAN	Q540F3 homo sapien
9	45	93.8	519	Q91NH0_9PARA	Q91NH0 tioman viru
10	44	91.7	147	Q6ZS83_HUMAN	Q6ZS83 homo sapien
11	44	91.7	252	Q54M51_DICDI	Q54M51 dictyosteli
12	44	91.7	355	Q5M9P4_RAT	Q5M9P4 rattus norv
13	43	89.6	162	Q43516_LYCSES	Q43516 lycopersico
14	43	89.6	162	Q6KA56_ORYSA	Q6KA56 oryza sativ
15	43	89.6	165	Q8BX70_MOUSE	Q8BX70 mus musculu
16	43	89.6	170	Q8H5L8_ORYSA	Q8H5L8 oryza sativ
17	43	89.6	209	1 CXX1_HUMAN	O15255 homo sapien
18	43	89.6	209	Q6IBF1_HUMAN	Q6IBF1 homo sapien
19	43	89.6	238	Q51C05_ENTHI	Q51C05 entamoeba h
20	43	89.6	243	Q4QHF1_LEIMA	Q4QHF1 leishmania
21	43	89.6	287	1 UDU3_ARATH	Q9FHD3 arabidopsis
22	43	89.6	462	2 Q9FB15_9ACTO	Q9FB15 streptomyce
23	43	89.6	462	2 Q9KW38_9RICK	Q9KW38 wolbachia s
24	43	89.6	486	2 Q731Y6_WOLPM	Q731Y6 wolbachia p
25	43	89.6	487	2 Q631B6_BURPS	Q631B6 burkholderi
26	43	89.6	488	2 Q4P068_USTWA	Q4P068 ustilago ma
27	43	89.6	493	2 Q52SK2_9RICK	Q52SK2 wolbachia e
28	43	89.6	551	2 Q4WXJ5_ASPFU	Q4WXJ5 aspergillus
29	43	89.6	554	2 Q6Z7E0_ORYSA	Q6Z7E0 oryza sativ
30	43	89.6	640	2 Q55P29_CRYNE	Q55P29 cryptococcus
31	43	89.6	745	2 Q6Z0N5_ORYSA	Q6Z0N5 oryza sativ

32	43	89.6	851	2 Q4T7J1_TETNG	Q4T7J1 tetraodon n
33	43	89.6	963	2 Q4SDM9_TETNG	Q4SDM9 tetraodon n
34	43	89.6	969	2 Q7RYU1_NEUCR	Q7RYU1 neurospora
35	43	89.6	1299	2 Q9UXA0_DROME	Q9UXA0 drosophila
36	43	89.6	1299	2 Q9V419_DROME	Q9V419 drosophila
37	43	89.6	2145	2 Q4QEP3_LEIMA	Q4QEP3 leishmania
38	43	89.6	2432	2 Q5XWJ3_9PICO	Q5XWJ3 aichi virus
39	43	89.6	2432	2 Q91QP4_9PICO	Q91QP4 aichi virus
40	43	89.6	2433	2 Q91464_9PICO	Q91464 aichi virus
41	43	89.6	2433	2 Q4TW10_9PICO	Q4TW10 aichi virus
42	42	87.5	80	2 Q5WLC4_BACSK	Q5WLC4 bacillus cl
43	42	87.5	106	2 Q4H8A2_9DEIO	Q4H8A2 deinococcus
44	42	87.5	134	2 Q5AG16_CANAL	Q5AG16 candida alb
45	42	87.5	144	2 Q5Z6V2_ORYSA	Q5Z6V2 oryza sativ
46	42	87.5	150	2 Q5ZB26_ORYSA	Q5ZB26 oryza sativ
47	42	87.5	151	2 Q8BGM6_MOUSE	Q8BGM6 m mus muscu
48	42	87.5	176	2 Q8PNK3_XANAC	Q8PNK3 xanthomonas
49	42	87.5	188	2 Q9D7T0_MOUSE	Q9D7T0 mus musculu
50	42	87.5	198	2 O01621_CABEL	O01621 caenorhabdi
51	42	87.5	280	1 TNFL6_GERTO	Q9BDN1 c tumor nec
52	42	87.5	280	1 TNFL6_MACFA	P63308 m tumor nec
53	42	87.5	280	1 TNFL6_MACMU	P63307 m tumor nec
54	42	87.5	280	1 TNFL6_MACNE	P63306 m tumor nec
55	42	87.5	280	2 Q861W5_FELCA	Q861W5 felis silve
56	42	87.5	281	1 TNFL6_HUMAN	P48023 h tumor nec
57	42	87.5	281	2 Q532Z1_HUMAN	Q532Z1 homo sapien
58	42	87.5	298	2 Q93F75_9BACI	Q93F75 bacillus sp
59	42	87.5	313	2 Q60D18_ORYSA	Q60D18 oryza sativ
60	42	87.5	355	2 Q84YW6_ORYSA	Q84YW6 oryza sativ
61	42	87.5	473	2 Q39620_CHLRE	Q39620 chlamydomon
62	42	87.5	503	2 Q9KW45_9RICK	Q9KW45 wolbachia s
63	42	87.5	519	2 Q7U110_RHOPA	Q7U110 rhodopirell
64	42	87.5	584	2 Q4RH23_TETNG	Q4RH23 tetraodon n
65	42	87.5	594	2 Q84Z97_ORYSA	Q84Z97 oryza sativ
66	42	87.5	606	2 Q852P0_VOLCA	Q852P0 voivox cart
67	42	87.5	643	2 Q8CB81_MOUSE	Q8CB81 mus musculu
68	42	87.5	665	2 Q80TL0_MOUSE	Q80TL0 mus musculu
69	42	87.5	749	2 Q5SX30_MOUSE	Q5SX30 mus musculu
70	42	87.5	750	2 Q80Z30_RAT	Q80Z30 rattus norv
71	42	87.5	757	2 Q7LAF3_HUMAN	Q7LAF3 homo sapien
72	42	87.5	759	2 Q9UPT0_HUMAN	Q9UPT0 homo sapien
73	42	87.5	766	2 Q8WY54_HUMAN	Q8WY54 homo sapien
74	42	87.5	773	2 Q4SHR4_TETNG	Q4SHR4 tetraodon n
75	42	87.5	837	2 Q5QMC9_ORYSA	Q5QMC9 oryza sativ
76	42	87.5	853	2 Q14333_HUMAN	Q14333 homo sapien
77	42	87.5	885	2 Q4RWK8_TETNG	Q4RWK8 tetraodon n
78	42	87.5	1035	2 Q5RIG8_BRACE	Q5RIG8 brachydanio
79	42	87.5	1077	2 Q5LCE8_BACFN	Q5LCE8 bacteroides
80	42	87.5	1077	2 Q64TJ3_BACFR	Q64TJ3 bacteroides
81	42	87.5	1157	2 Q87217_NEUCR	Q87217 neurospora
82	42	87.5	1272	2 Q54PT8_DICDI	Q54PT8 dictyosteli
83	42	87.5	1430	1 CSK11_RAT	Q8VHK2 rattus norv
84	42	87.5	2775	2 Q5B770_EMENI	Q5B770 aspergillus
85	41	85.4	40	2 Q62KF6_BURMA	Q62KF6 burkholderi
86	41	85.4	101	2 Q8Q870_9BETA	Q8Q870 pongine her
87	41	85.4	128	2 Q8STT3_ENCCU	Q8STT3 encephalito
88	41	85.4	128	2 Q67TS6_ORYSA	Q67TS6 oryza sativ
89	41	85.4	130	2 Q6YUJ6_ORYSA	Q6YUJ6 oryza sativ
90	41	85.4	131	2 Q4UR12_XANCP	Q4UR12 xanthomonas
91	41	85.4	131	2 Q8PC16_XANCP	Q8PC16 xanthomonas
92	41	85.4	151	2 Q9F195_ARATH	Q9F195 arabidopsis
93	41	85.4	156	1 VFG_PLRV	P10471 potato leaf
94	41	85.4	156	2 Q84828_PLRV	Q84828 potato leaf
95	41	85.4	156	2 Q8QYN5_PLRV	Q8QYN5 potato leaf
96	41	85.4	156	2 Q8QYQ5_PLRV	Q8QYQ5 potato leaf
97	41	85.4	156	2 Q9QGS2_PLRV	Q9QGS2 potato leaf
98	41	85.4	188	2 Q67TQ3_ORYSA	Q67TQ3 oryza sativ
99	41	85.4	199	2 Q7XR59_ORYSA	Q7XR59 oryza sativ
100	41	85.4	240	2 Q4HY80_GIBZE	Q4HY80 gibberella
101	41	85.4	260	2 Q82HP0_STRAW	Q82HP0 streptomyce
102	41	85.4	271	2 Q80WJ0_MOUSE	Q80WJ0 mus musculu
103	41	85.4	279	2 Q6N908_RHOPA	Q6N908 rhodopseudo
104	41	85.4	294	2 Q8W0Q8_SORBI	Q8W0Q8 sorghum bic

105	41	85.4	320	2	Q6Y27_ORYSA	Q6Y27_ORYSA	178	41	85.4	1487	2	Q68E09_HUMAN	Q68E09 homo sapien
106	41	85.4	321	2	Q9HCN7_HUMAN	Q9HCN7 homo sapien	179	41	85.4	1600	2	Q4I646_GIBBEZ	Q4I646 gibberella
107	41	85.4	337	2	Q5FPY8_GLUOX	Q5FPY8 gluconobact	180	41	85.4	1651	2	Q6ZML7_HUMAN	Q6ZML7 homo sapien
108	41	85.4	339	2	Q9UIF2_HUMAN	Q9UIF2 homo sapien	181	41	85.4	1837	2	Q9VCZ9_DROME	Q9VCZ9 drosophila
109	41	85.4	354	2	Q4I2G8_GIBZE	Q4I2G8 gibberella	182	41	85.4	2054	2	Q7QAT6_ANOVA	Q7QAT6 anopheles g
110	41	85.4	355	2	Q4ZVK0_PSESV	Q4ZVK0 pseudomonas	183	40	83.3	111	1	Y3835_GLOVI	Y3835 gloeobacter
111	41	85.4	378	2	Q5YME9_NOCPA	Q5YME9 nocardia fa	184	40	83.3	111	2	Q9WAZ9_9VIRU	Q9WAZ9 torque teno
112	41	85.4	386	2	Q1I714_CABEL	Q1I714 caenorhabdi	185	40	83.3	111	2	Q9WB00_9VIRU	Q9WB00 torque teno
113	41	85.4	401	2	Q6PA50_XENLA	Q6PA50 xenopus lae	186	40	83.3	111	2	Q9WB03_9VIRU	Q9WB03 torque teno
114	41	85.4	404	2	Q6DP84_XENLA	Q6DP84 xenopus lae	187	40	83.3	112	2	Q9QU23_9VIRU	Q9QU23 torque teno
115	41	85.4	423	2	Q87X22_NEUCR	Q87X22 neurospora	188	40	83.3	121	2	Q9DTD3_9VIRU	Q9DTD3 torque teno
116	41	85.4	431	2	Q8IRH4_DROME	Q8IRH4 drosophila	189	40	83.3	128	2	Q8V7C9_9VIRU	Q8V7C9 streptomyc
117	41	85.4	443	2	Q6VVV7_ORYSA	Q6VVV7 rattus sativ	190	40	83.3	134	2	Q9X8L8_STRCO	Q9X8L8 streptomyc
118	41	85.4	459	1	MIG6_RAT	Q05432 rattus norv	191	40	83.3	156	1	VPG_PLRV1	P17524 potato leaf
119	41	85.4	461	1	MIG6_MOUSE	Q09J27 mus musculu	192	40	83.3	156	1	VPG_PLRV1	P17523 potato leaf
120	41	85.4	462	1	MIG6_MOUSE	Q09J27 mus musculu	193	40	83.3	156	1	VPG_PLRV1	P17523 potato leaf
121	41	85.4	477	2	Q9NSG8_HUMAN	Q9NSG8 homo sapien	194	40	83.3	156	2	Q7E888_PLRV	Q7E888 potato leaf
122	41	85.4	488	2	Q5IAX29_ENTHI	Q5IAX29 entamoeba h	195	40	83.3	156	2	Q84839_PLRV	Q84839 potato leaf
123	41	85.4	495	2	Q9H517_HUMAN	Q9H517 homo sapien	196	40	83.3	156	2	Q8QYP2_PLRV	Q8QYP2 potato leaf
124	41	85.4	499	2	Q5IAX34_ENTHI	Q5IAX34 entamoeba h	197	40	83.3	156	2	Q8QYP8_PLRV	Q8QYP8 potato leaf
125	41	85.4	526	2	Q50IWI_RAT	Q50IWI rattus norv	198	40	83.3	156	2	Q8QYR2_PLRV	Q8QYR2 potato leaf
126	41	85.4	545	1	FGR_FSVGR	P00544 feline sarc	199	40	83.3	156	2	Q9GGS5_PLRV	Q9GGS5 potato leaf
127	41	85.4	548	1	PTP61_DROME	Q9W091 drosophila	200	40	83.3	156	2	Q9YNY3_PLRV	Q9YNY3 potato leaf
128	41	85.4	566	1	UBP21_MOUSE	Q9QZ16 mus musculu	201	40	83.3	156	2	Q70805_9VIRU	Q70805 torque teno
129	41	85.4	586	2	Q9DSX3_MOUSE	Q9DSX3 mus musculu	202	40	83.3	156	2	Q9WSM8_9VIRU	Q9WSM8 torque teno
130	41	85.4	601	1	KCNAS_MUSPF	P79197 mustela put	203	40	83.3	157	2	Q9JH34_9VIRU	Q9JH34 torque teno
131	41	85.4	627	2	Q9XVH7_CABEL	Q9XVH7 caenorhabdi	204	40	83.3	165	2	Q64M82_ORYSA	Q64M82 oryza sativ
132	41	85.4	628	2	Q4YQ7_9DELT	Q4YQ7 anaeromyxob	205	40	83.3	172	2	Q9DS34_MOUSE	Q9DS34 mus musculu
133	41	85.4	631	2	Q4XOH0_ASFPF	Q4XOH0 aspergillus	206	40	83.3	173	2	Q6ZOM9_ORYSA	Q6ZOM9 oryza sativ
134	41	85.4	631	2	Q5XIL5_RAT	Q5XIL5 rattus norv	207	40	83.3	204	2	Q9WSM6_9VIRU	Q9WSM6 torque teno
135	41	85.4	641	1	MAL12_HUMAN	Q60476 homo sapien	208	40	83.3	205	2	Q655L3_ORYSA	Q655L3 oryza sativ
136	41	85.4	658	2	Q66HC8_RAT	Q66HC8 rattus norv	209	40	83.3	215	2	Q18661_CABEL	Q18661 caenorhabdi
137	41	85.4	664	2	Q4RTR5_TETNG	Q4RTR5 tetraodon n	210	40	83.3	238	2	Q8GR34_DESVU	Q8GR34 desulfovibr
138	41	85.4	673	2	Q5EBP4_MOUSE	Q5EBP4 mus musculu	211	40	83.3	244	2	Q63NAB_BURPS	Q63NAB burkholderi
139	41	85.4	675	2	Q8QWJ1_MOUSE	Q8QWJ1 mus musculu	212	40	83.3	245	2	O61572_CABEL	O61572 caenorhabdi
140	41	85.4	678	2	Q7RWV5_NEUCR	Q7RWV5 neurospora	213	40	83.3	247	2	Q4Q3G7_THRPA	Q4Q3G7 thelleria a
141	41	85.4	682	2	Q4IIQ6_GIBZE	Q4IIQ6 gibberella	214	40	83.3	247	2	Q4Q3G7_THRPA	Q4Q3G7 thelleria a
142	41	85.4	740	2	Q59U10_CANAL	Q59U10 candida alb	215	40	83.3	273	2	O61670_MYCTU	O61670 mycobacteri
143	41	85.4	742	2	Q59U34_CANAL	Q59U34 candida alb	216	40	83.3	273	2	O7TVF3_MYCTU	O7TVF3 mycobacteri
144	41	85.4	753	2	Q6PIJ3_HUMAN	Q6PIJ3 homo sapien	217	40	83.3	275	2	Q7D707_MYCTU	Q7D707 mycobacteri
145	41	85.4	760	2	Q8NBU7_HUMAN	Q8NBU7 homo sapien	218	40	83.3	283	2	Q6ZJL2_ORYSA	Q6ZJL2 oryza sativ
146	41	85.4	774	1	GLIS3_HUMAN	Q8NEA6 homo sapien	219	40	83.3	310	2	O5YX54_NOCPA	O5YX54 nocardia fa
147	41	85.4	775	2	Q5VZV9_HUMAN	Q5VZV9 homo sapien	220	40	83.3	382	2	Q6Q047_CABER	Q6Q047 caenorhabdi
148	41	85.4	775	2	Q4SME7_TETNG	Q4SME7 tetraodon n	221	40	83.3	382	2	Q9GYM4_CABEL	Q9GYM4 caenorhabdi
149	41	85.4	780	2	Q8LX78_ARATH	Q8LX78 arabidopsis	222	40	83.3	385	2	Q6K649_ORYSA	Q6K649 oryza sativ
150	41	85.4	781	2	Q823Y3_CHLCV	Q823Y3 chlamydophi	223	40	83.3	421	2	O5SKB6_THET8	O5SKB6 thermus the
151	41	85.4	782	2	Q8GX37_ARATH	Q8GX37 arabidopsis	224	40	83.3	421	2	Q72KM8_THET2	Q72KM8 thermus the
152	41	85.4	812	2	Q9DMG7_RCMVM	Q9DMG7 rat cymomeg	225	40	83.3	429	2	Q75M69_ORYSA	Q75M69 oryza sativ
153	41	85.4	824	2	Q5QKN9_ORYSA	Q5QKN9 oryza sativ	226	40	83.3	488	2	O5Z1J6_NOCPA	O5Z1J6 nocardia fa
154	41	85.4	847	2	Q6UXZ5_HUMAN	Q6UXZ5 homo sapien	227	40	83.3	501	2	O5Z5Z3_ORYSA	O5Z5Z3 oryza sativ
155	41	85.4	847	2	Q5R9H3_PONPY	O5R9H3 pongo pygma	228	40	83.3	527	2	O5JXK0_ORYSA	O5JXK0 oryza sativ
156	41	85.4	849	2	Q7T299_BRABE	Q7T299 brachydanio	229	40	83.3	530	2	O5KC56_CRYNE	O5KC56 cryptococcu
157	41	85.4	874	2	Q4UPW5_THRPA	Q4UPW5 thelleria a	230	40	83.3	546	2	O5IMH6_MAGGR	O5IMH6 magnaporth
158	41	85.4	894	2	O22824_ARATH	O22824 arabidopsis	231	40	83.3	553	2	Q6DN78_LOLPR	Q6DN78 lolium pere
159	41	85.4	934	2	Q8SXX4_DROME	Q8SXX4 drosophila	232	40	83.3	564	2	O9LDA9_ARATH	O9LDA9 arabidopsis
160	41	85.4	1022	2	Q4RWM9_TETNG	Q4RWM9 tetraodon n	233	40	83.3	566	2	O86854_STRCO	O86854 streptomyc
161	41	85.4	1054	2	Q4QOW2_LEIMA	Q4QOW2 leishmania	234	40	83.3	574	1	DPOA2_SCHPO	DPOA2 schizosacch
162	41	85.4	1125	2	O5CHA3_CRYHO	Q4QW2 cryptocoppi	235	40	83.3	574	2	Q69B22_ASPPA	Q69B22 aspergillus
163	41	85.4	1151	2	Q4S8G7_TETNG	Q4S8G7 tetraodon n	236	40	83.3	577	2	Q75UH4_DICDI	Q75UH4 dictyostell
164	41	85.4	1204	2	Q8IXB2_HUMAN	Q8IXB2 homo sapien	237	40	83.3	601	2	Q4I5K8_GIBZE	Q4I5K8 gibberella
165	41	85.4	1204	2	Q53EP0_HUMAN	Q53EP0 homo sapien	238	40	83.3	629	2	Q4FX88_LEIMA	Q4FX88 leishmania
166	41	85.4	1204	2	Q4SOT1_TETNG	Q4SOT1 homo sapien	239	40	83.3	748	2	O5IWM9_MAGGR	O5IWM9 magnaporth
167	41	85.4	1219	2	Q4SOT1_TETNG	Q4SOT1 tetraodon n	240	40	83.3	765	2	O55UM0_CRYNE	O55UM0 cryptococcu
168	41	85.4	1241	2	Q9W3L1_DROME	Q9W3L1 drosophila	241	40	83.3	765	2	O5XHS4_CRYNE	O5XHS4 cryptococcu
169	41	85.4	1292	2	Q20330_CABEL	Q20330 caenorhabdi	242	40	83.3	775	2	Q9US20_CHLPN	Q9US20 chlamydia p
170	41	85.4	1353	2	Q55UG1_CRYNE	Q55UG1 cryptococcu	243	40	83.3	775	2	Q9Z879_CHLPN	Q9Z879 chlamydia p
171	41	85.4	1353	2	O5KHY3_CRYNE	O5KHY3 cryptococcu	244	40	83.3	790	2	Q82LN3_STRAW	Q82LN3 streptomyc
172	41	85.4	1359	2	Q7RXQ6_NEUCR	Q7RXQ6 neurospora	245	40	83.3	791	1	SP3A1_MOUSE	SP3A1 mouse
173	41	85.4	1363	2	Q4STY2_TETNG	Q4STY2 tetraodon n	246	40	83.3	791	2	O5ZM84_CHICK	O5ZM84 gallus gall
174	41	85.4	1370	2	Q7SOP7_NEUCR	Q7SOP7 neurospora	247	40	83.3	793	1	SP3A1_HUMAN	SP3A1 human
175	41	85.4	1379	2	O5AZX0_EMENI	O5AZX0 aspergillus	248	40	83.3	797	2	Q9RKS9_STRCO	Q9RKS9 streptomyc
176	41	85.4	1409	2	Q4XIX2_ASFPF	Q4XIX2 aspergillus	249	40	83.3	809	2	Q8RWQ1_ARATH	Q8RWQ1 arabidopsis
177	41	85.4	1473	2	Q4FVZ1_LEIMA	Q4FVZ1 leishmania	250	40	83.3	809	2	Q8CAB2_MOUSE	Q8CAB2 mus musculu

251	40	83.3	864	1	ADA15_MOUSE	O88839	mus musculus	39	81.2	675	2	Q6UEI2_PNECA	Q6uei2_pneumocysti
252	40	83.3	864	2	Q6P779_RAT	Q6p779	rattus norv	39	81.2	749	2	Q57VY5_9TRYP	Q57vy5_trypanosoma
253	40	83.3	929	2	Q59P09_CANAL	Q59p09	candida alb	326	81.2	753	2	Q6ZIP3_ORYSA	Q6zip3_oryza sativ
254	40	83.3	944	2	Q4Q979_LEIMA	Q4q979	leishmania	327	81.2	780	2	Q9K402_STRCO	Q9k402_streptomyce
255	40	83.3	1008	2	Q72KA4_THET2	Q72ka4	thermus the	328	81.2	788	2	Q69MT2_ORYSA	Q69mt2_oryza sativ
256	40	83.3	1008	2	Q5SJY3_THET8	Q5sjv3	thermus the	329	81.2	809	2	O76869_DROME	O76869_drosophila
257	40	83.3	1048	2	Q5AYK5_EMENI	Q5ayk5	aspergillus	330	81.2	824	2	O4RID3_TETNG	O4rid3_tetradon n
258	40	83.3	1367	2	Q552E9_DICDI	Q552e9	dictyosteli	331	81.2	845	2	O6ZH84_ORYSA	O6zh84_oryza sativ
259	40	83.3	1424	2	Q5XJY6_MOUSE	Q5xjv6	mus musculus	332	81.2	851	1	DDX31_HUMAN	O9w4v1_drosophila
260	40	83.3	1424	2	Q52KFI_MOUSE	Q52kfi	mus musculus	333	81.2	853	2	O4IAH9_GIBZE	O4iah9_gibberella
261	40	83.3	2371	2	O58NA5_CHLIN	O58na5	chlamydomon	334	81.2	859	2	O8IRV3_DROME	O8irv3_drosophila
262	40	83.3	3889	2	Q6SSP8_CHLRE	Q6sse8	chlamydomon	335	81.2	861	2	O55XX6_CRYNE	O55xx6_cryptococcu
263	40	83.3	3930	2	Q54HU6_DICDI	Q54hj6	dictyosteli	336	81.2	861	2	O5KM19_CRYNE	O5km19_cryptococcu
264	39	81.2	117	2	Q8RYA1_ORYSA	O8ryal	oryza sativ	337	81.2	872	2	O6ZA49_ORYSA	O6zaa9_oryza sativ
265	39	81.2	127	2	Q55KG3_MACFA	Q55kg3	macaca fasc	338	81.2	887	2	O9W4V1_DROME	O9w4v1_drosophila
266	39	81.2	130	2	Q8WZJ3_9APHY	Q8wzj3	dictyonema	339	81.2	923	2	O4I4L5_GIBZE	O4i4l5_gibberella
267	39	81.2	130	2	O5NN73_ZYMMO	O5nn73	zymononas m	340	81.2	939	2	O4PG76_USTMA	O4pg76_ustilago ma
268	39	81.2	132	2	Q57WK9_9TRYP	Q57wk9	trypanosoma	341	81.2	948	2	O4TKS9_9SPHN	O4tkx9_erythrobact
269	39	81.2	132	2	Q57Y16_9TRYP	Q57yl6	trypanosoma	342	81.2	962	2	O61PF1_CAEBR	O61pfi_caenorhabdi
270	39	81.2	134	2	Q5DUH9_9VIRU	Q5duh9	torque teno	343	81.2	1006	2	O750H0_ASHGO	O750h0_ashbya goss
271	39	81.2	134	2	Q5DUH9_9VIRU	Q5duh9	torque teno	343	81.2	1006	2	O750H0_ASHGO	O750h0_ashbya goss
272	39	81.2	137	2	O8IYK6_HUMAN	O8iyl6	homo sapien	344	81.2	1104	2	O55HE0_CRYNE	O55he0_cryptococcu
273	39	81.2	153	1	PRIMA_HUMAN	O86xrs	homo sapien	345	81.2	1104	2	O5K6X9_CRYNE	O5k6x9_cryptococcu
274	39	81.2	153	1	PRIMA_MOUSE	O81of0	mus musculus	346	81.2	1106	2	O16310_CAEEL	O16310_caenorhabdi
275	39	81.2	153	2	Q5TFG9_HUMAN	Q5tfgr	homo sapien	347	81.2	1141	2	O7S421_NEUCR	O7s421_neurospora
276	39	81.2	153	2	Q51542_PSEAE	Q51542	pseudomonas	348	81.2	1211	2	O35233_MOUSE	O35233_mus musculus
277	39	81.2	172	2	Q6Z4V3_ORYSA	Q6z4v3	oryza sativ	349	81.2	1217	1	APFI_MOUSE	O88573_mus musculus
278	39	81.2	173	2	O6Z063_CAEEL	Q6z063	caenorhabdi	350	81.2	1218	2	O5XM34_MOUSE	O5xm34_mus musculus
279	39	81.2	189	2	O84QR5_ORYSA	Q84qr5	oryza sativ	351	81.2	1233	2	O7SCT1_NEUCR	O7sct1_neurospora
280	39	81.2	216	2	O4NPG8_9DELT	Q4npg8	anaeromyxob	352	81.2	1285	2	O8K3T3_SPAJD	O8k3t3_spalax juda
281	39	81.2	251	2	Q6IJW2_DROME	Q6ijw2	drosophila	353	81.2	1318	2	O55LL3_CRYNE	O55ll3_cryptococcu
282	39	81.2	252	2	Q5Z732_ORYSA	Q5z732	oryza sativ	354	81.2	1318	2	O5KD53_CRYNE	O5kd53_cryptococcu
283	39	81.2	307	2	Q9VZ28_DROME	Q9vz28	drosophila	355	81.2	1426	2	O4RTA6_TETNG	O4rta6_tetradon n
284	39	81.2	309	2	Q8UZB4_9VIRU	Q8uzb4	grapevine f	356	81.2	1462	2	O4TBB7_TETNG	O4tbb7_tetradon n
285	39	81.2	318	2	O4I545_WHEAT	Q4i545	triticum ae	357	81.2	1698	2	O4RZU0_TETNG	O4rzu0_tetradon n
286	39	81.2	323	2	O7V622_PROMM	Q7v622	prochloroco	358	81.2	1733	1	VNUA_PRYKA	P33485_pseudorabie
287	39	81.2	344	1	TILS_SYNEL	Q8dih2	synechococc	359	81.2	1733	2	O5PPC8_9ALPH	O5ppc8_suid herpes
288	39	81.2	355	1	SIRT6_HUMAN	Q8n6t7	homo sapien	360	81.2	1795	2	O746C5_THET2	O746c5_thermus the
289	39	81.2	373	2	O7TXG5_MYCBO	Q7txg5	mycobacteri	361	81.2	1795	2	O53VY7_THET8	O53vy7_thermus the
290	39	81.2	373	2	O53253_MYCTU	Q53253	mycobacteri	362	81.2	1878	2	O66580_EEEV	O66580_eastern equ
291	39	81.2	378	1	PRX4_MOUSE	Q59ne9	mus musculus	363	81.2	1951	2	O6VZM1_CNVP	O6vzml_canarypox v
292	39	81.2	409	1	YNP2_YEAST	P53901	saccharomyc	364	81.2	1958	2	O69340_9ALPH	O69340_suid herpes
293	39	81.2	409	2	Q6O576_YEAST	Q6g5j6	saccharomyc	365	81.2	2156	2	O7SGT6_NEUCR	O7sgt6_neurospora
294	39	81.2	423	2	Q4NS24_9DELT	Q4ns24	anaeromyxob	366	81.2	2493	2	O88789_EEEV	O88789_eastern equ
295	39	81.2	424	2	O6FP92_CANGA	Q6fp92	candida gla	367	81.2	2494	2	O5XZP4_EEEV	O5xzf4_eastern equ
296	39	81.2	434	2	O69TP3_ORYSA	Q69tp3	oryza sativ	368	81.2	2494	2	O4QXJ8_EEEV	O4qxj8_eastern equ
297	39	81.2	439	2	Q4NX07_9DELT	Q4nx07	anaeromyxob	369	81.2	2494	2	O4QXK0_EEEV	O4qxk0_eastern equ
298	39	81.2	440	1	BCL3_MOUSE	Q9z2f6	mus musculus	370	81.2	2511	2	O95LC8_DIDMA	O95lc8_didelphis m
299	39	81.2	444	2	Q9C197_AAMU	Q9c197	amanita mus	371	81.2	2615	2	O9CPI1_PASMU	O9cpil1_pasteurella
300	39	81.2	446	1	BCL3_HUMAN	P20749	homo sapien	372	81.2	2705	2	O9W6V6_CHICK	O9w6v6_gallus gall
301	39	81.2	459	2	O8GMG2_STRGL	Q8rgm2	streptomyce	373	81.2	2725	2	O9UKZ4_HUMAN	O9ukz4_homo sapien
302	39	81.2	472	2	O9RVZ5_DEIRA	Q9rvz5	deinococcus	374	81.2	2725	2	O5JZ17_HUMAN	O5jz17_homo sapien
303	39	81.2	477	2	Q6ZW60_HUMAN	Q6zw60	homo sapien	375	81.2	2731	2	O9WTS4_MOUSE	O9wts4_mus musculus
304	39	81.2	487	2	O51KP3_MAGGR	Q51kp3	magnaporthe	376	81.2	2757	2	O5CPF7_CRYPV	O5cpf7_cryptospori
305	39	81.2	493	2	Q9LT85_ARATH	Q9lt85	arabidopsis	377	81.2	3027	2	O6UDW5_PLAFA	O6udw5_plasmodium
306	39	81.2	502	2	Q7R7A8_PLAYO	Q7r7a8	plasmodium	378	81.2	3596	1	RPOA_SHFV	O68772_s_replicase
307	39	81.2	506	2	Q504F6_BRARE	Q504f6	brachydanio	379	81.2	76	2	O7S0D9_NEUCR	O7s0d9_neurospora
308	39	81.2	519	2	O91MK3_9PARA	Q91mk3	menangle vi	380	81.2	105	2	O7EYTA_ORYSA	O7eyta_oryza sativ
309	39	81.2	523	2	O93XJ2_PERFR	Q93xj2	perilla fru	381	81.2	115	2	O7ZFB7_DESVH	O7zfb7_desulfovibr
310	39	81.2	532	2	O6ZD77_ORYSA	Q6zd77	oryza sativ	382	81.2	117	2	O6TV80_ORYSA	O6tv80_oryza sativ
311	39	81.2	533	2	O86ZC3_USTMA	Q86zc3	ustilago ma	383	81.2	121	1	AMEL_ORNAN	O97646_ornithorhyn
312	39	81.2	533	2	O4PC32_USTMA	Q4pc32	ustilago ma	384	81.2	121	1	AMEL_TACAC	O97647_tachyglossus
313	39	81.2	536	2	O8C8D2_MOUSE	Q8c8d2	mus musculus	385	81.2	135	2	O4UUE9_XANCP	O4uee9_xanthomonas
314	39	81.2	538	2	O6C0T9_YARLI	Q6c0t9	yarrowia li	386	81.2	135	2	O8P9E8_XANCP	O8p9e8_xanthomonas
315	39	81.2	561	1	EFS_HUMAN	O43281	homo sapien	387	81.2	139	2	O9BE61_MACFA	O9be61_macaca fasc
316	39	81.2	580	2	Q7SEW0_NEUCR	Q7sew0	neurospora	388	81.2	141	2	O51E04_ENTHI	O51e04_entamoeba h
317	39	81.2	593	2	Q28248_CANPA	Q28248	canis famil	389	81.2	154	2	O5KNJ7_CRYNE	O5knj7_cryptococcu
318	39	81.2	600	2	Q866G0_CANFA	Q866g0	canis famil	390	81.2	155	2	O7XW07_ORYSA	O7xw07_oryza sativ
319	39	81.2	602	1	KCNAS_MOUSE	Q86162	mus musculus	391	81.2	156	2	O8ZRS7_STRAW	O8zrs7_streptomyce
320	39	81.2	602	2	Q9Z1R6_MOUSE	Q9z1r6	mus musculus	392	81.2	178	2	O7XHQ5_ORYSA	O7xhq5_oryza sativ
321	39	81.2	609	2	Q4RFZ2_TETNG	Q4rfz2	tetradon n	393	81.2	178	2	O98680_9BETA	O98680_simian cyto
322	39	81.2	612	2	O503F8_BRARE	Q503f8	brachydanio	394	81.2	181	2	O4LYX8_9BURK	O4lyx8_burkholderi
323	39	81.2	616	2	O42200_BRARE	O42200	brachydanio	395	81.2	191	2	O9A416_CAUCR	O9a416_caulobacter
324	39	81.2	636	2	O4SF15_NEUCR	Q4sf15	neurospora	396	81.2	192	2	O4I3J4_GIBZE	O4i3j4_gibberella

397	38	79.2	192	2	Q53LW2_ORYSA	Q53LW2_ORYSA sativ	470	38	79.2	475	2	Q5RE17_PONPY	Q5RE17_pongo pygma
398	38	79.2	194	2	Q7PPI4_ANOGA	Q7PPI4 anopheles g	471	38	79.2	478	2	Q54LJ3_DICDI	Q54LJ3 dictyosteli
399	38	79.2	200	2	Q8LAT8_ARATH	Q8LAT8 arabidopsis	472	38	79.2	484	2	Q5YQH1_NOCFA	Q5YQH1 noctuidia fa
400	38	79.2	204	2	Q9LT71_ARATH	Q9LT71 arabidopsis	473	38	79.2	485	1	T13H_TAXCU	T13H taxus cuspi
401	38	79.2	204	2	Q72H82_THET2	Q72H82 thermus the	474	38	79.2	485	2	Q5BU48_9CONI	Q5BU48 taxus x med
402	38	79.2	204	2	Q5MD8_THET8	Q5MD8 thermus the	475	38	79.2	485	2	Q56GD5_TAXCH	Q56GD5 taxus chine
403	38	79.2	207	2	Q98679_9BETA	Q98679 simlan cyto	476	38	79.2	487	1	WASIP_RAT	WASIP rattus norv
404	38	79.2	208	2	Q6QI49_RAT	Q6QI49 rattus norv	477	38	79.2	489	2	Q4SFL7_TETNG	Q4SFL7 tetraodon n
405	38	79.2	216	2	Q623X0_ORYSA	Q623X0 oryza sativ	478	38	79.2	491	2	Q67V14_ORYSA	Q67V14 oryza sativ
406	38	79.2	217	2	Q5N952_ORYSA	Q5N952 oryza sativ	479	38	79.2	495	2	Q6JD68_TAXCA	Q6JD68 taxus canad
407	38	79.2	217	2	Q5N952_ORYSA	Q5N952 oryza sativ	480	38	79.2	495	2	Q5SIU2_TAXCH	Q5SIU2 taxus chine
408	38	79.2	217	2	Q9M4Y5_ORYSA	Q9M4Y5 oryza sativ	481	38	79.2	514	2	Q62418_CAEBR	Q62418 caenorhabdi
409	38	79.2	222	2	Q51B85_ENTHI	Q51B85 entamoeba h	482	38	79.2	527	1	TXK_HUMAN	TXK mus musculu
410	38	79.2	222	2	Q51B85_ENTHI	Q51B85 entamoeba h	483	38	79.2	527	1	TXK_MOUSE	TXK mus musculu
411	38	79.2	229	2	Q6K9N1_DEIRA	Q6K9N1 deinococcus	484	38	79.2	533	2	Q75IR3_ORYSA	Q75IR3 oryza sativ
412	38	79.2	230	2	Q6K9N3_ORYSA	Q6K9N3 oryza sativ	485	38	79.2	542	2	Q5LIU7_MAGGR	Q5LIU7 magnaporth
413	38	79.2	230	2	Q41979_MHV68	Q41979 murid herpe	486	38	79.2	544	2	Q6H082_PREDI	Q6H082 fremyella d
414	38	79.2	238	2	Q67YV6_ORYSA	Q67YV6 oryza sativ	487	38	79.2	560	2	Q8IYU1_HUMAN	Q8IYU1 homo sapien
415	38	79.2	240	2	Q521N1_MAGGR	Q521N1 magnaporth	488	38	79.2	561	2	Q8G4W0_BIFLO	Q8G4W0 bifidobacte
416	38	79.2	247	2	Q6CSL7_YARLI	Q6CSL7 yarrowia li	489	38	79.2	570	2	Q55SI1_CRYNE	Q55SI1 cryptococcu
417	38	79.2	256	2	Q6Z074_ORYSA	Q6Z074 oryza sativ	490	38	79.2	585	2	Q5H067_XANOR	Q5H067 xanthomonas
418	38	79.2	262	2	Q6ZRP0_HUMAN	Q6ZRP0 homo sapien	491	38	79.2	587	2	Q7S6M3_NEUCR	Q7S6M3 neurospora
419	38	79.2	262	2	Q6X263_GALPH	Q6X263 bovine herp	492	38	79.2	590	2	Q6CDQ2_YARLI	Q6CDQ2 yarrowia li
420	38	79.2	266	2	Q41HX5_GIBZE	Q41HX5 gibberella	493	38	79.2	599	2	Q68FAS_XENTR	Q68FAS xenopus tro
421	38	79.2	271	2	Q6D047_ERWCT	Q6D047 erwinia car	494	38	79.2	602	2	Q7SGU0_NEUCR	Q7SGU0 neurospora
422	38	79.2	280	2	Q5S1L9_THET8	Q5S1L9 thermus the	495	38	79.2	605	1	PP2CD_HUMAN	PP2CD homo sapien
423	38	79.2	280	2	Q721Z6_THET2	Q721Z6 thermus the	496	38	79.2	605	2	Q53XP4_HUMAN	Q53XP4 homo sapien
424	38	79.2	282	2	Q7XGJ1_ORYSA	Q7XGJ1 oryza sativ	497	38	79.2	605	2	Q8NEA7_HUMAN	Q8NEA7 homo sapien
425	38	79.2	282	2	Q94HM7_ORYSA	Q94HM7 oryza sativ	498	38	79.2	605	2	Q6P991_HUMAN	Q6P991 homo sapien
426	38	79.2	287	2	Q7Y092_ORYSA	Q7Y092 oryza sativ	499	38	79.2	605	2	Q7SYD2_ERARE	Q7SYD2 brachydanio
427	38	79.2	292	2	Q4WSM5_ASFPF	Q4WSM5 aspergillus	500	38	79.2	612	2	Q4SN07_TETNG	Q4SN07 tetraodon n
428	38	79.2	300	2	Q6MYZ3_ASFPF	Q6MYZ3 aspergillus	501	38	79.2	624	2	Q8A69_XYFT	Q8A69 xylella fas
429	38	79.2	302	2	Q4NP7_9DELT	Q4NP7 anaeromyxob	502	38	79.2	636	2	Q6I21_EPTBU	Q6I21 eptaretus
430	38	79.2	303	2	Q7YV35_9TRYP	Q7YV35 trypanosoma	503	38	79.2	639	1	ZIC5_HUMAN	ZIC5 homo sapien
431	38	79.2	304	2	Q876G4_SACHA	Q876G4 saccharomyc	504	38	79.2	639	2	Q5VYB0_HUMAN	Q5VYB0 homo sapien
432	38	79.2	305	2	Q6YWH3_ORYSA	Q6YWH3 oryza sativ	505	38	79.2	644	2	Q7RIV9_PLAYO	Q7RIV9 plaemophilu
433	38	79.2	306	2	Q4RUA7_TETNG	Q4RUA7 tetraodon n	506	38	79.2	649	2	Q9UIM6_DROME	Q9UIM6 drosophila
434	38	79.2	308	1	MPASA_IOLPR	Q40240 lolium pere	507	38	79.2	653	2	Q6UPR4_CHLRE	Q6UPR4 chlamydomon
435	38	79.2	318	1	ADT2_YEAST	P18239 saccharomyc	508	38	79.2	656	2	Q7S6U6_NEUCR	Q7S6U6 neurospora
436	38	79.2	324	2	Q4STL5_TETNG	Q4STL5 tetraodon n	509	38	79.2	657	1	BTX_CHICK	BTX gallus gall
437	38	79.2	330	2	Q5QNI5_ORYSA	Q5QNI5 oryza sativ	510	38	79.2	657	2	Q5ZLF2_CHICK	Q5ZLF2 gallus gall
438	38	79.2	331	2	Q8NRR3_CORGL	Q8NRR3 corynebacte	511	38	79.2	665	2	Q5KGX2_CRYNE	Q5KGX2 cryptococcu
439	38	79.2	336	2	Q84BF5_STRLI	Q84BF5 streptomyce	512	38	79.2	665	2	Q5OXG4_ENTHI	Q5OXG4 entamoeba h
440	38	79.2	339	2	Q524X6_MAGGR	Q524X6 magnaporth	513	38	79.2	668	1	PBS2_YEAST	PBS2 saccharomyc
441	38	79.2	344	2	Q5T8S7_HUMAN	Q5T8S7 homo sapien	514	38	79.2	677	2	Q6FL74_CANGA	Q6FL74 candida gla
442	38	79.2	349	2	Q87X32_PSESM	Q87X32 pseudomonas	515	38	79.2	681	2	Q6BUS0_DEBHA	Q6BUS0 debaryomyc
443	38	79.2	356	2	Q924W1_RAT	Q924W1 rattus norv	516	38	79.2	683	2	Q96W50_DEBHA	Q96W50 debaryomyc
444	38	79.2	360	2	Q719H0_BRAFL	Q719H0 branchiosto	517	38	79.2	687	2	Q7QDV6_ANOGA	Q7QDV6 anopheles g
445	38	79.2	364	2	Q5SQI6_CRYNE	Q5SQI6 cryptococcu	518	38	79.2	687	2	Q5IG68_ENTHI	Q5IG68 entamoeba h
446	38	79.2	364	2	Q5KFN0_CRYNE	Q5KFN0 cryptococcu	519	38	79.2	691	2	Q7Y5N1_ASHGO	Q7Y5N1 ashbya goss
447	38	79.2	374	2	Q746C6_THET2	Q746C6 thermus the	520	38	79.2	705	2	Q8CHU0_NEUCR	Q8CHU0 neurospora
448	38	79.2	374	2	Q7V8X4_BORPE	Q7V8X4 bordetella	521	38	79.2	707	2	Q8CHU0_MOUSE	Q8CHU0 mus musculu
449	38	79.2	374	2	Q7W2T8_BORPA	Q7W2T8 bordetella	522	38	79.2	708	2	Q8SX98_DROME	Q8SX98 drosophila
450	38	79.2	387	2	Q84HB1_STRCZ	Q84HB1 streptomyc	523	38	79.2	724	2	Q6CN49_KJULA	Q6CN49 kiuyveromyc
451	38	79.2	398	2	Q8VBA1_WSSV	Q8VBA1 white spot	524	38	79.2	738	2	Q92213_RAT	Q92213 rattus norv
452	38	79.2	398	2	Q911L1_WSSV	Q911L1 white spot	525	38	79.2	740	2	Q5TTF1_ANOGA	Q5TTF1 anopheles g
453	38	79.2	400	2	Q8FOS2_COREF	Q8FOS2 corynebacte	526	38	79.2	751	2	Q8MLU0_DROME	Q8MLU0 drosophila
454	38	79.2	402	2	Q4RX44_TETNG	Q4RX44 tetraodon n	527	38	79.2	753	2	Q7KVL6_DROME	Q7KVL6 drosophila
455	38	79.2	403	2	Q4SEW0_TETNG	Q4SEW0 tetraodon n	528	38	79.2	761	2	Q4RRE9_TETNG	Q4RRE9 tetraodon n
456	38	79.2	404	2	Q8X095_NEUCR	Q8X095 neurospora	529	38	79.2	774	2	Q63X79_BURPS	Q63X79 burkholderi
457	38	79.2	407	2	Q744V4_MYCPA	Q744V4 mycobacteri	530	38	79.2	787	2	Q62MM5_BURMA	Q62MM5 burkholderi
458	38	79.2	407	2	Q9AA59_CAUCR	Q9AA59 caulobacter	531	38	79.2	837	2	Q60Z34_CASBR	Q60Z34 caenorhabdi
459	38	79.2	408	1	YHFW_ECOLI	P45549 escherichia	532	38	79.2	870	2	Q7VZJ7_BORPE	Q7VZJ7 bordetella
460	38	79.2	408	2	Q8FCW4_ECOL6	Q8FCW4 escherichia	533	38	79.2	870	2	Q7WHV9_BORBR	Q7WHV9 bordetella
461	38	79.2	408	2	Q83PX2_SHIFL	Q83PX2 shigella fl	534	38	79.2	874	2	Q5JU58_HUMAN	Q5JU58 homo sapien
462	38	79.2	418	2	Q9KYZ8_STRCO	Q9KYZ8 streptomyc	535	38	79.2	876	2	Q6XLI5_BRARE	Q6XLI5 brachydanio
463	38	79.2	425	2	Q9C0Q7_NEUCR	Q9C0Q7 neurospora	536	38	79.2	883	2	Q4RUA9_TETNG	Q4RUA9 tetraodon n
464	38	79.2	427	2	Q5ASC8_EMENI	Q5ASC8 aspergillus	537	38	79.2	884	2	Q5NBA4_ORYSA	Q5NBA4 oryza sativ
465	38	79.2	430	2	Q8IVR6_HUMAN	Q8IVR6 homo sapien	538	38	79.2	885	2	Q4PHI1_USMTA	Q4PHI1 ustilago ma
466	38	79.2	439	2	Q61989_CAEBR	Q61989 caenorhabdi	539	38	79.2	888	2	Q5B371_EMENI	Q5B371 aspergillus
467	38	79.2	443	1	SAM68_MOUSE	Q60749 m kh domain	540	38	79.2	894	2	Q7W9W8_BORPA	Q7W9W8 bordetella
468	38	79.2	443	1	SAM68_MOUSE	Q91V33 r kh domain	541	38	79.2	901	2	Q53VY8_THET8	Q53VY8 thermus the
469	38	79.2	474	2	Q5B433_EMENI	Q5B433 aspergillus	542	38	79.2				

543 38 79.2 922 2 QBUVP2_COTCO
544 38 79.2 925 1 PIP1_YEAST
545 38 79.2 925 2 QAIRH1_GIBZE
546 38 79.2 949 2 QSUJX1_HUMAN
547 38 79.2 995 1 MA2B2_PIG
548 38 79.2 1003 2 Q9C235_NEUCR
549 38 79.2 1006 2 Q55R68_CRYNE
550 38 79.2 1006 2 Q5KFO2_CRYNE
551 38 79.2 1009 2 Q55R69_CRYNE
552 38 79.2 1009 2 Q5KFO1_CRYNE
553 38 79.2 1009 2 Q4CEU8_LEIMA
554 38 79.2 1057 2 Q4CES9_YARLI
555 38 79.2 1101 2 Q5K9C7_CRYNE
556 38 79.2 1101 2 Q55K74_CRYNE
557 38 79.2 1125 2 Q5BPF7_EMENI
558 38 79.2 1134 2 Q80T51_MOUSE
559 38 79.2 1157 1 DHX37_HUMAN
560 38 79.2 1169 2 Q5KI84_CRYNE
561 38 79.2 1169 2 Q55T11_CRYNE
562 38 79.2 1218 2 Q9V6T7_DROME
563 38 79.2 1229 2 Q4MS24_ASPFU
564 38 79.2 1236 2 Q8MY63_ASPFU
565 38 79.2 1379 2 Q8I7P4_DICDI
566 38 79.2 1379 2 Q54D97_DICDI
567 38 79.2 1393 2 Q55U57_CRYNE
568 38 79.2 1398 2 Q7F1V5_ORYSA
569 38 79.2 1402 2 Q5KI87_CRYNE
570 38 79.2 1478 2 Q5JU85_HUMAN
571 38 79.2 1560 2 Q60275_HUMAN
572 38 79.2 1677 2 Q4RV09_TETNG
573 38 79.2 1700 2 Q9VJ39_DROME
574 38 79.2 1914 2 Q8X010_NEUCR
575 38 79.2 1919 2 Q7RW47_NEUCR
576 38 79.2 2025 2 Q4X1U1_ASPFU
577 38 79.2 2039 2 Q7S3G9_NEUCR
578 38 79.2 2161 1 SHAN1_HUMAN
579 38 79.2 2295 2 Q7S2P2_NEUCR
580 38 79.2 2837 2 Q96U29_NEUCR
581 38 79.2 3980 2 Q4T089_TETNG
582 37.5 78.1 71 2 Q6MRN1_BDEBA
583 38 77.1 52 2 Q6SPR5_SHEEP
584 37 77.1 55 2 Q9TUJ6_CERNI
585 37 77.1 60 2 Q6SPR6_CEREL
586 37 77.1 75 2 Q4SHG6_TETNG
587 37 77.1 79 2 Q9PX71_9BETA
588 37 77.1 83 2 Q4GYV7_TETNG
589 37 77.1 93 2 Q58N36_9CAUD
590 37 77.1 99 2 Q5SMJ8_ORYSA
591 37 77.1 104 2 Q9YBH3_AERPE
592 37 77.1 104 2 Q7IRB1_HUMAN
593 37 77.1 106 2 Q693N3_CLOAL
594 37 77.1 107 2 Q7EZ89_ORYSA
595 37 77.1 109 2 Q9SY21_ARATH
596 37 77.1 123 2 Q8BNW7_MOUSE
597 37 77.1 124 2 Q82UK1_RHIME
598 37 77.1 126 2 Q67VU5_ORYSA
599 37 77.1 129 2 Q72I10_THET2
600 37 77.1 143 2 Q7PNB6_ANOGA
601 37 77.1 151 2 Q84PV1_ORYSA
602 37 77.1 151 2 Q4NA22_9HICC
603 37 77.1 162 2 Q6ZP66_HUMAN
604 37 77.1 166 2 Q9X8W3_STRCO
605 37 77.1 167 2 Q4PMF1_IXOSC
606 37 77.1 175 2 Q83739_9LUTE
607 37 77.1 181 2 Q6SP07O_AZOSE
608 37 77.1 189 2 Q8Z6J1_ORYSA
609 37 77.1 194 2 Q7NCAS_GLOVI
610 37 77.1 196 2 Q69TW5_ORYSA
611 37 77.1 202 2 Q75KQ8_ORYSA
612 37 77.1 203 2 Q9CKX9_STRCO
613 37 77.1 206 2 Q7NUT6_CHRVO
614 37 77.1 210 2 Q3VEH6_DROME
615 37 77.1 211 2 Q55UJ5_CRYNE

616 37 77.1 211 2 Q5KHV0_CRYNE
617 37 77.1 211 2 Q51H25_ENTHI
618 37 77.1 220 2 Q60EX0_ORYSA
619 37 77.1 220 2 Q6Z915_ORYSA
620 37 77.1 223 2 Q55TL7_CRYNE
621 37 77.1 223 2 Q5KIR8_CRYNE
622 37 77.1 228 2 Q60EB3_ORYSA
623 37 77.1 232 2 Q6Y2N8_ORYSA
624 37 77.1 237 2 Q6AW70_VIRU
625 37 77.1 244 2 Q5YSD4_NOCPA
626 37 77.1 245 2 Q63LF2_BURFS
627 37 77.1 250 2 Q7ZYI9_XENLA
628 37 77.1 254 2 Q6U536_9HEPC
629 37 77.1 254 2 Q6U541_9HEPC
630 37 77.1 256 2 Q71B87_BRUAB
631 37 77.1 258 2 Q32585_ECOLI
632 37 77.1 258 2 Q8GJ13_ECOLI
633 37 77.1 263 2 Q5GX23_XANOR
634 37 77.1 263 2 Q5GX32_XANOR
635 37 77.1 265 2 Q52415_MAGGR
636 37 77.1 267 2 Q9BHP8_LEIMA
637 37 77.1 271 2 Q67W47_ORYSA
638 37 77.1 272 2 Q8BUJ5_MOUSE
639 37 77.1 273 2 Q82HN1_STRAW
640 37 77.1 276 2 Q72U85_LEPIC
641 37 77.1 276 2 Q8F0V1_LEPIN
642 37 77.1 279 2 Q4TPV6_9SPHN
643 37 77.1 280 2 Q5QLI3_ORYSA
644 37 77.1 281 2 Q6Z6E2_ORYSA
645 37 77.1 285 2 P78977_YARLI
646 37 77.1 286 2 Q67TR5_ORYSA
647 37 77.1 287 2 Q4WUC2_ASPFU
648 37 77.1 287 2 Q67WE9_ORYSA
649 37 77.1 287 2 Q5DTS1_MOUSE
650 37 77.1 289 2 Q5Z572_ORYSA
651 37 77.1 289 2 Q7EYB7_ORYSA
652 37 77.1 297 2 Q9SXE7_ARATH
653 37 77.1 298 2 Q9RVV7_DEIRA
654 37 77.1 299 2 Q93F63_9BACI
655 37 77.1 304 2 Q4LRL5_PETHY
656 37 77.1 304 2 Q4QV52_9STLI
657 37 77.1 313 2 Q7S7C4_NEUCR
658 37 77.1 314 2 Q54817_MOUSE
659 37 77.1 317 2 Q6MW23_NEUCR
660 37 77.1 318 2 Q4ZLH4_FSESY
661 37 77.1 321 2 Q53SB7_HUMAN
662 37 77.1 322 2 Q70K98_9ACTO
663 37 77.1 322 2 Q9IK80_NPVST
664 37 77.1 323 2 Q6MES3_PARUW
665 37 77.1 343 2 Q8LNI0_ORYSA
666 37 77.1 343 2 Q4ZLH3_PSESY
667 37 77.1 343 2 Q8D8U2_VIBU
668 37 77.1 346 2 Q7NGK0_GLOVI
669 37 77.1 346 2 Q7MLN2_VIBV
670 37 77.1 348 2 Q4S820_TETNG
671 37 77.1 349 2 Q4ZP39_PSESY
672 37 77.1 355 2 Q55L60_CRYNE
673 37 77.1 355 2 Q5KAC1_CRYNE
674 37 77.1 357 2 Q9VJ16_DROME
675 37 77.1 362 2 Q4TPE3_TETNG
676 37 77.1 368 2 Q7XG36_ORYSA
677 37 77.1 368 2 Q94120_ORYSA
678 37 77.1 370 2 Q827M8_STRAW
679 37 77.1 372 2 Q529G0_MAGGR
680 37 77.1 375 2 Q9NPS0_HUMAN
681 37 77.1 376 2 Q9H7E6_HUMAN
682 37 77.1 376 2 Q4Q177_LEIMA
683 37 77.1 382 2 Q88933_MOUSE
684 37 77.1 383 2 Q7YIE1_ORYSA
685 37 77.1 386 2 Q9FPQ5_CHLRE
686 37 77.1 387 2 Q922U3_MOUSE
687 37 77.1 392 2 Q4R9F1_MACFA
688 37 77.1 397 2 Q8K437_MOUSE

Q8UVF2 coturnix co
P40020 saccharomyc
Q4IRH1 gibberella
Q8JUX1 homo sapien
Q88949 sus scrofa
Q9C235 neurospora
Q55R68 cryptococcu
Q5KFO2 cryptococcu
Q55R69 cryptococcu
Q5KFO1 cryptococcu
Q4CEU8 leishmania
Q4CES9 yarrowia li
Q5K9C7 cryptococcu
Q55K74 cryptococcu
Q5BPF7 aspergillus
Q80T51 mus musculu
Q80T37 homo sapien
Q5KI84 cryptococcu
Q55T11 cryptococcu
Q9V6T7 drosophila
Q4W824 aspergillus
Q8MY63 aspergillus
Q8I7P4 dictyosteli
Q54D97 dictyosteli
Q55U57 cryptococcu
Q7F1V5 oryza sativ
Q5KI87 cryptococcu
Q5JU85 homo sapien
Q60275 homo sapien
Q4RV09 tetradodon n
Q9VJ39 drosophila
Q8X010 neurospora
Q7RW47 neurospora
Q4X1U1 aspergillus
Q7S3G9 neurospora
Q9Y566 homo sapien
Q7S2P2 neurospora
Q96U29 neurospora
Q4T089 tetradodon n
Q6MRN1 bellovibri
Q6SPR5 ovis aries
Q9TUJ6 cervus nipp
Q6SPR6 cervus elap
Q4SHG6 tetradodon n
Q9PX71 human herpe
Q4GYV7 trypanosoma
Q58N36 cyanophaga
Q5SMJ8 oryza sativ
Q9YBH3 aeropyrum p
Q7IRB1 homo sapien
Q693N3 clogmia alb
Q7EZ89 oryza sativ
Q9SY21 arabidopsis
Q8BNW7 mus musculu
Q82UK1 rhizobium m
Q67VU5 oryza sativ
Q72I10 thermus the
Q7PNB6 anopheles g
Q84PV1 oryza sativ
Q4NA22 arthrobacte
Q6ZP66 homo sapien
Q9X8W3 streptomyce
Q4PMF1 ixodes scap
Q83739 beet mild y
Q93739 azoarcus sp
Q5P07O oryza sativ
Q8Z6J1 gloeobacter
Q7NCAS gloeobacter
Q69TW5 oryza sativ
Q75KQ8 oryza sativ
Q9CKX9 streptomyce
Q7NUT6 chromobacte
Q3VEH6 drosophila
Q55UJ5 cryptococcu

Q5KHV0 cryptococcu
Q51H25 entamoeba h
Q60EX0 oryza sativ
Q6Z915 oryza sativ
Q55TL7 cryptococcu
Q5KIR8 cryptococcu
Q60EB3 oryza sativ
Q6Y2N8 oryza sativ
Q6AW70 bombyx mori
Q5YSD4 nocardia fa
Q63LF2 burkholderi
Q7ZYI9 xenopus lae
Q6U536 hepatitis c
Q6U541 hepatitis c
Q71B87 brucella ab
Q32585 escherichia
Q8GJ13 escherichia
Q5GX23 xanthomonas
Q5GX32 xanthomonas
Q52415 magnaporthe
Q9BHP8 leishmania
Q67W47 oryza sativ
Q8BUJ5 mus musculu
Q82HN1 streptomyce
Q72U85 leptospira
Q8F0V1 leptospira
Q4TPV6 erythrobaet
Q5QLI3 oryza sativ
Q6Z6E2 oryza sativ
P78977 yarrowia li
Q67TR5 oryza sativ
Q4WUC2 aspergillus
Q67WE9 oryza sativ
Q5DTS1 mus musculu
Q5Z572 oryza sativ
Q7EYB7 oryza sativ
Q9SXE7 arabidopsis
Q9RVV7 deinococcus
Q93F63 bacillus sp
Q4LRL5 petunia t-ly
Q4QV52 simian t-ly
Q7S7C4 mus musculu
Q54817 mus musculu
Q6MW23 neurospora
Q4ZLH4 pseudomonas
Q53SB7 homo sapien
Q70K98 gordonia we
Q9IK80 spodoptera
Q6MES3 parachlamyd
Q8LNI0 oryza sativ
Q4ZLH3 pseudomonas
Q8D8U2 vibrio vuln
Q7NGK0 gloeobacter
Q7MLN2 vibrio vuln
Q4S820 tetradodon n
Q4ZP39 pseudomonas
Q55L60 cryptococcu
Q5KAC1 cryptococcu
Q9VJ16 drosophila
Q4TPE3 tetradodon n
Q7XG36 oryza sativ
Q94120 oryza sativ
Q827M8 streptomyce
Q529G0 magnaporthe
Q9NPS0 homo sapien
Q9H7E6 homo sapien
Q4Q177 leishmania
Q88933 mus musculu
Q7YIE1 oryza sativ
Q9FPQ5 chlamydomon
Q922U3 mus musculu
Q4R9F1 macaca faec
Q8K437 mus musculu

589	37	77.1	397	2	Q8K0J2_MOUSE	Q8k0j2 mus musculus	762	37	77.1	663	1	TERM_ADEB3	O55439 bovine aden
590	37	77.1	397	2	Q6PA02_MOUSE	Q6pa02 mus musculus	763	37	77.1	665	2	Q8XR57_RALSO	O8xr57 ralsconia s
591	37	77.1	400	2	Q73VH2_MYCPA	Q73vh2 mycobacteri	764	37	77.1	669	2	O8H8J8_ORYSA	O8h8j8 oryza sativ
592	37	77.1	408	2	O44776_CABEL	O44776 caenorhabdi	765	37	77.1	676	1	EXTI1_HUMAN	O92935 homo sapien
593	37	77.1	409	2	Q9SBM1_VOLCA	Q9sbm1 volvox cart	766	37	77.1	676	2	Q6GSC1_HUMAN	O6gsc1 homo sapien
594	37	77.1	411	2	Q7S8I9_NEUCR	Q7s8i9 neurospora	767	37	77.1	686	2	Q5B4D7_EMENI	O5b4d7 aspergillus
595	37	77.1	413	2	O15417_HUMAN	O15417 homo sapien	768	37	77.1	688	2	O87I13_NEUCR	O87i13 neurospora
596	37	77.1	414	2	O55LS8_CRYNE	O55ls8 cryptococcu	769	37	77.1	691	2	O55S15_CRYNE	O55s15 cryptococcu
597	37	77.1	414	2	O5K827_CRYNE	O5k827 cryptococcu	770	37	77.1	691	2	O5KGX4_CRYNE	O5kgx4 cryptococcu
598	37	77.1	414	2	O8NBX7_HUMAN	O8nbx7 homo sapien	771	37	77.1	693	2	O4P3N9_USTMA	O4p3n9 ustilago ma
599	37	77.1	417	2	O6PIN9_HUMAN	O6pin9 homo sapien	772	37	77.1	695	2	Q9XUQ0_CABEL	Q9xuq0 caenorhabdi
700	37	77.1	426	2	O5YU88_NOCFA	O5yu88 nocardia fa	773	37	77.1	697	2	O6LK69_CABBR	O6lk69 caenorhabdi
701	37	77.1	429	2	O5BFA3_EMENI	O5bfa3 aspergillus	774	37	77.1	700	2	O8N0R5_BOMMO	O8n0r5 bombyx mori
702	37	77.1	430	2	O5TQ09_ANOGA	O5tq09 anopheles g	775	37	77.1	703	2	O9BYI8_HUMAN	O9byi8 homo sapien
703	37	77.1	435	1	RAD52_ASHGO	Q756f4 ashbya goss	776	37	77.1	703	2	O8N6Y0_HUMAN	O8n6y0 homo sapien
704	37	77.1	445	2	O95KJ7_MACFA	Q95kj7 macaca fasc	777	37	77.1	724	2	O8BWD1_MOUSE	O8bwd1 mus musculus
705	37	77.1	446	2	O5TTT0_ANOGA	O5ttt0 anopheles g	778	37	77.1	745	2	O9LFT3_ARATH	O9lft3 arabidopsis
706	37	77.1	454	2	Q9RVH3_DEIRA	Q9rvh3 deinococcus	779	37	77.1	752	2	Q90ZP4_ONCHY	Q90zp4 oncorhynch
707	37	77.1	471	2	O9NPM5_HUMAN	O9npms homo sapien	780	37	77.1	757	2	Q4NEX0_9DELT	Q4nrx0 anaeromykob
708	37	77.1	481	2	Q82CH7_STRAW	Q82ch7 streptomyce	781	37	77.1	760	2	Q66BK5_YERPS	Q66bk5 yersinia ps
709	37	77.1	485	2	Q7UA85_SYNPK	Q7ua85 synecococc	782	37	77.1	767	2	Q4ZM3_9DELT	Q4zm3 anaeromykob
710	37	77.1	487	1	EBN2_EBV	P12978 Epstein-Bar	783	37	77.1	776	2	O6K3K2_ORYSA	O6k3k2 oryza sativ
711	37	77.1	487	2	Q777H1_9GAMA	Q777h1 human herpe	784	37	77.1	777	2	Q7S7K1_NEUCR	Q7s7k1 neurospora
712	37	77.1	488	2	Q6CBT3_YARLI	Q6cbt3 yarrowia li	785	37	77.1	780	2	Q4NEV6_ASFPF	Q4nev6 aspergillus
713	37	77.1	488	2	Q95UR9_SCHAM	Q95ur9 schistocerc	786	37	77.1	782	2	Q4WEV6_ASFPF	Q4wev6 aspergillus
714	37	77.1	488	2	Q4Q2C0_LEIMA	Q4q2c0 leishmania	787	37	77.1	784	2	O53S01_HUMAN	O53s01 homo sapien
715	37	77.1	489	2	O873K2_NEUCR	O873k2 neurospora	788	37	77.1	784	2	O86NK9_DROME	O86nk9 drosophila
716	37	77.1	489	2	Q73XK9_MYCPA	Q73xk9 mycobacteri	789	37	77.1	786	2	O5TTK2_ANOGA	O5ttk2 anopheles g
717	37	77.1	490	2	Q690Z3_9GAMA	Q690z3 human herpe	790	37	77.1	794	2	Q7RYS4_NEUCR	Q7rys4 neurospora
718	37	77.1	499	2	Q6YTE2_ORYSA	Q6yte2 oryza sativ	791	37	77.1	821	2	O5KCM4_CRYNE	O5kcm4 cryptococcu
719	37	77.1	500	2	Q69X58_ORYSA	Q69x58 oryza sativ	792	37	77.1	825	2	O6GN93_XENLA	O6gn93 xenopus lae
720	37	77.1	500	2	O6YTF1_ORYSA	O6ytf1 oryza sativ	793	37	77.1	830	2	O95UR8_SCHAM	O95ur8 schistocerc
721	37	77.1	514	2	O8X0Q3_NEUCR	O8x0q3 neurospora	794	37	77.1	836	2	Q4RVH6_TETNG	Q4rvh6 tetraodon n
722	37	77.1	519	2	O4I9W3_GIBZE	O4i9w3 gibberella	795	37	77.1	838	2	O8R3S1_MOUSE	O8r3s1 mus musculus
723	37	77.1	520	2	Q8R4X9_MOUSE	Q8r4x9 mus musculus	796	37	77.1	845	2	Q7S4M4_NEUCR	Q7s4m4 neurospora
724	37	77.1	520	2	Q99MT6_MOUSE	Q99mt6 mus musculus	797	37	77.1	851	2	O66BA9_YERPS	O66ba9 yersinia ps
725	37	77.1	521	2	Q73XJ0_MYCPA	Q73xj0 mycobacteri	798	37	77.1	863	2	Q9LEE8_MAIZE	Q9lee8 zea mays (m
726	37	77.1	521	2	Q99P73_MOUSE	Q99p73 mus musculus	799	37	77.1	867	2	O4X0Y7_ASFPF	O4x0y7 aspergillus
727	37	77.1	521	2	O8VDL9_MOUSE	O8vdl9 mus musculus	800	37	77.1	871	2	O6DJB9_XENTR	O6djb9 xenopus tro
728	37	77.1	524	2	O9B2E0_HUMAN	O9bze0 homo sapien	801	37	77.1	876	2	O9P4Y9_NEUCR	O9p4y9 neurospora
729	37	77.1	535	2	O55Y03_CRYNE	O55y03 cryptococcu	802	37	77.1	880	1	PHIT_DROME	Q9v9a8 drosophila
730	37	77.1	535	2	O5KLY7_CRYNE	O5kly7 cryptococcu	803	37	77.1	880	2	O614D5_CABBR	O614d5 caenorhabdi
731	37	77.1	542	2	O55J32_CRYNE	O55j32 cryptococcu	804	37	77.1	885	2	O5SKS5_CRYNE	O5ska5 cryptococcu
732	37	77.1	547	2	O9CA82_ARATH	O9ca82 arabidopsis	805	37	77.1	885	2	O5KAQ0_CRYNE	O5kaq0 cryptococcu
733	37	77.1	551	2	O9NP74_HUMAN	O9np74 homo sapien	806	37	77.1	908	2	O80VR4_MOUSE	O80vr4 mus musculus
734	37	77.1	555	1	GPI_CHLRE	O9f06 chlamydomon	807	37	77.1	916	2	O59LR9_CANAL	O59lr9 candida alb
735	37	77.1	563	2	O613F0_CABBR	O613f0 caenorhabdi	808	37	77.1	936	2	O8QRV7_9BETA	O8qrv7 pongine her
736	37	77.1	564	2	Q7SCU8_NEUCR	Q7scu8 neurospora	809	37	77.1	939	2	Q4SPD1_TETNG	Q4spd1 tetraodon n
737	37	77.1	569	2	Q9CU24_MOUSE	Q9cu24 mus musculus	810	37	77.1	943	2	Q7SDY9_NEUCR	Q7sdy9 neurospora
738	37	77.1	580	2	O51UJ2_MAGGR	O51uj2 magnaporth	811	37	77.1	951	2	Q7RUP0_NEUCR	Q7rup0 neurospora
739	37	77.1	582	1	MNT_HUMAN	O99583 homo sapien	812	37	77.1	970	2	O60L67_CABBR	O60l67 caenorhabdi
740	37	77.1	582	2	O6MT9_NEUCR	O6mvt9 neurospora	813	37	77.1	971	1	R3HDM_HUMAN	Q15032 homo sapien
741	37	77.1	582	2	O5NSV7_SYNPF	O5nsv7 synecococc	814	37	77.1	1005	2	O8MQS4_CABEL	O8mq54 caenorhabdi
742	37	77.1	589	2	Q8RW91_ARATH	O8rw91 arabidopsis	815	37	77.1	1011	2	O9P944_PNECA	O9p944 pneumocysti
743	37	77.1	591	1	MNT_MOUSE	O08789 mus musculus	816	37	77.1	1014	2	O6DRC5_BRARE	O6drc5 brachydanio
744	37	77.1	591	2	Q6GTJ3_MOUSE	O6gtj3 mus musculus	817	37	77.1	1019	2	O4RG15_TETNG	O4rg15 tetraodon n
745	37	77.1	593	2	O9MA04_ARATH	O9ma04 arabidopsis	818	37	77.1	1024	2	O4XWM8_ASFPF	O4xwm8 aspergillus
746	37	77.1	599	2	O80VL9_MOUSE	O80vl9 mus musculus	819	37	77.1	1024	2	Q9XGS7_CHLRE	O9xgs7 chlamydomon
747	37	77.1	604	2	Q9L252_STRCO	Q9l252 streptomyce	820	37	77.1	1061	2	O97XK5_PONPY	O97xk5 pongo pygma
748	37	77.1	608	2	Q4HZQ7_CIOIN	Q4hzq7 ciona intes	821	37	77.1	1062	2	Q4Q9P4_LETMA	Q4q9p4 leishmania
749	37	77.1	610	2	O6ZC88_ORYSA	O6zc88 oryza sativ	822	37	77.1	1063	2	O4V809_RUBV	O4v809 rubella vir
750	37	77.1	612	2	O8F9J9_CORRF	O8f9j9 corynebacte	823	37	77.1	1069	2	O5B2T2_MAGGR	O5b2t2 magnaporth
751	37	77.1	616	2	Q8T0J3_DROME	O8t0j3 drosophila	824	37	77.1	1099	2	O81W32_HUMAN	O81w32 homo sapien
752	37	77.1	633	2	O5Z5T3_ORYSA	O5z5t3 oryza sativ	825	37	77.1	1116	2	Q7PLF4_DROME	Q7plf4 drosophila
753	37	77.1	640	2	Q60W81_CABBR	Q60w81 caenorhabdi	826	37	77.1	1126	2	O64115_XENLA	O64115 xenopus lae
754	37	77.1	645	2	Q7RW18_NEUCR	Q7rw18 neurospora	827	37	77.1	1140	2	O20299_CABEL	O20299 caenorhabdi
755	37	77.1	646	1	FBX46_HUMAN	O6pj61 homo sapien	828	37	77.1	1146	2	O6PLP6_CHLRE	O6plp6 chlamydomon
756	37	77.1	647	2	Q96FT7_HUMAN	O96ft7 homo sapien	829	37	77.1	1147	1	SPR15_HUMAN	O95104 homo sapien
757	37	77.1	648	2	O5GYL8_XANOR	O5gyl8 xanthomonas	830	37	77.1	1147	2	O5D4Z9_9EURO	O5d4z9 monascus pi
758	37	77.1	648	2	O4UTQ5_XANCP	O4utq5 xanthomonas	831	37	77.1	1147	2	O5D500_9EURO	O5d500 monascus pi
759	37	77.1	648	2	Q8P9I0_XANCP	Q8p9i0 xanthomonas	832	37	77.1	1154	2	Q5HZU5_XANOR	O5hzus xanthomonas
760	37	77.1	648	2	O8PLQ7_XANAC	O8plq7 xanthomonas	833	37	77.1	1154	2	O8PIN7_XANAC	O8pin7 xanthomonas
761	37	77.1	657	2	Q7N7J1_PHOLL	Q7n7j1 photorhabdu	834	37	77.1	1155	2	Q4UWT9_XANCP	Q4uwt9 xanthomonas

835	37	77.1	1155	2	QBP7B3_XANCP	QBP7b3_xanthomonas	908	36	75.0	87	2	Q69MY3_ORYSA	Q69my3_oryza sativ
836	37	77.1	1183	2	Q7TS6_MAGR	Q7ts6_mus musculus	909	36	75.0	88	2	Q5Z714_ORYSA	Q5z714_oryza sativ
837	37	77.1	1191	2	Q51UB0_MAGR	Q51ub0_magnaporthe	910	36	75.0	89	2	Q7WXL8_ALCEU	Q7wxl8_alcaligenes
838	37	77.1	1200	2	Q69ZP8_MOUSE	Q69zp8_mus musculus	911	36	75.0	93	2	Q6UUI3_ORYSA	Q6uui3_oryza sativ
839	37	77.1	1201	2	Q24240_DROME	Q24240_drosophila	912	36	75.0	101	2	Q8C877_MOUSE	Q8c877_mus musculus
840	37	77.1	1209	2	Q6PPF0_MOUSE	Q6ppf0_mus musculus	913	36	75.0	102	2	Q5JNC8_ORYSA	Q5jnc8_oryza sativ
841	37	77.1	1210	2	Q753G0_ASHGO	Q753g0_ashbya goss	914	36	75.0	103	2	Q84JC7_ORYSA	Q84jc7_oryza sativ
842	37	77.1	1212	2	Q8MK2_DROME	Q8mk2_drosophila	915	36	75.0	104	1	PK25_PEA	PK1892_plaum sativ
843	37	77.1	1216	2	Q4S2D6_TETNG	Q4s2d6_tetradodon n	916	36	75.0	104	2	Q9FW1_ORYSA	Q9fw1_oryza sativ
844	37	77.1	1239	2	Q9V4P1_DROME	Q9v4p1_drosophila	917	36	75.0	105	2	Q9D7W2_MOUSE	Q9d7w2_mus musculus
845	37	77.1	1262	1	STNB_DROME	Q24212_drosophila	918	36	75.0	106	2	Q9FKP0_ARATH	Q9fkp0_arabidopsis
846	37	77.1	1297	2	Q9V604_DROME	Q9v604_drosophila	919	36	75.0	107	2	Q8LEL5_ARATH	Q8lel5_arabidopsis
847	37	77.1	1325	2	Q96PW1_HUMAN	Q96pw1_homo sapien	920	36	75.0	108	2	Q6A688_PROAC	Q6a688_propionibac
848	37	77.1	1375	2	Q4MZV7_ASPFU	Q4mzv7_aspergillus	921	36	75.0	111	2	Q749V8_GEOSL	Q749v8_geobacter s
849	37	77.1	1391	2	Q6K8Z4_ORYSA	Q6k8z4_oryza sativ	922	36	75.0	112	2	Q7EYJ0_ORYSA	Q7eyj0_oryza sativ
850	37	77.1	1408	2	Q6X125_9ALPH	Q6x125_bovine herp	923	36	75.0	112	2	Q88EX3_FSEPK	Q88ex3_pseudomonas
851	37	77.1	1426	1	BCL9_HUMAN	Q00512_homo sapien	924	36	75.0	114	2	Q4PEW6_USTMA	Q4pew6_ustilago ma
852	37	77.1	1426	2	Q5T4S9_HUMAN	Q5t4s9_homo sapien	925	36	75.0	116	2	P82411_SPIOL	P82411_spinacia ol
853	37	77.1	1434	1	SYNJ2_MOUSE	Q3d2g5_mus musculus	926	36	75.0	116	2	Q8CDE5_MOUSE	Q8cde5_mus musculus
854	37	77.1	1464	2	Q96782_DROME	Q96782_drosophila	927	36	75.0	117	2	Q57Y88_9TRYP	Q57y88_trypanosoma
855	37	77.1	1464	2	Q9VX75_DROME	Q9vx75_drosophila	928	36	75.0	120	2	Q53783_STRAM	Q53783_streptomyce
856	37	77.1	1495	2	Q5KIY9_CRYNE	Q5kiy9_cryptococcu	929	36	75.0	121	2	Q49510_ARATH	Q49510_arabidopsis
857	37	77.1	1496	1	SYNJ2_RAT	Q55207_rattus norv	930	36	75.0	123	2	Q8UGH2_AGR75	Q8ugh2_agrobacteri
858	37	77.1	1525	2	Q5ST89_CRYNE	Q5st89_cryptococcu	931	36	75.0	124	2	Q96PJ0_HUMAN	Q96pj0_homo sapien
859	37	77.1	1553	2	Q4P0W7_USTMA	Q4p0w7_ustilago ma	932	36	75.0	124	2	Q7Z2T9_HUMAN	Q7z2t9_homo sapien
860	37	77.1	1560	2	Q6JF2_HUMAN	Q6jp2_homo sapien	933	36	75.0	124	2	Q6YD3_ORYSA	Q6yjd3_oryza sativ
861	37	77.1	1581	2	Q6BU20_DEBHA	Q6bu20_debaryomyce	934	36	75.0	126	2	Q69X78_ORYSA	Q69x78_oryza sativ
862	37	77.1	1589	2	Q9UO9_METAN	Q9uog9_metharizium	935	36	75.0	128	2	Q6QNH4_9ORTH	Q6qnh4_chorthippus
863	37	77.1	1661	2	Q5TQ89_ANOGA	Q5tqe9_anopheles g	936	36	75.0	129	2	Q5PQ17_XENLA	Q5pq17_xenopus lae
864	37	77.1	1704	2	Q41200_GIBZE	Q412j0_gibberella	937	36	75.0	129	2	Q4TGN5_TETNG	Q4tgn5_tetradodon n
865	37	77.1	1758	2	Q16625_CAEEL	Q16625_caenorhabdi	938	36	75.0	133	2	Q950M7_9FUNG	Q950m7_rhizophydia
866	37	77.1	1790	2	Q84214_BRARE	Q84214_brachydanio	939	36	75.0	135	2	Q6YSJ4_ORYSA	Q6ysj4_oryza sativ
867	37	77.1	1884	2	Q5CS67_CRYPV	Q5cs67_cryptospori	940	36	75.0	136	2	Q8SYT0_DROME	Q8syto_drosophila
868	37	77.1	1922	2	Q5CID8_CRYHO	Q5cid8_cryptospori	941	36	75.0	136	2	Q61IM7_DROME	Q61im7_drosophila
869	37	77.1	2103	2	Q5K4K9_CHLRE	Q5kan9_chlamydomon	942	36	75.0	138	2	Q87115_NEUCR	Q87115_neurospora
870	37	77.1	2375	2	Q5AN08_CANAL	Q5an08_candida alb	943	36	75.0	138	2	Q5QN36_ORYSA	Q5qn36_oryza sativ
871	37	77.1	2383	2	Q5AMK3_CANAL	Q5amk3_candida alb	944	36	75.0	139	2	Q7X190_ORYSA	Q7xi90_oryza sativ
872	37	77.1	2472	1	NCOR2_MOUSE	Q9wu42_mus musculus	945	36	75.0	140	2	Q827D8_STRAW	Q827d8_streptomyce
873	37	77.1	2551	2	Q4P0X4_USTMA	Q4p0x4_ustilago ma	946	36	75.0	141	2	Q4N6K2_THEPA	Q4n6k2_theileria p
874	37	77.1	2594	2	Q4SVP5_TETNG	Q4svd5_tetradodon n	947	36	75.0	142	2	Q4J3P2_AZOVI	Q4j3p2_azotobacter
875	37	77.1	2710	2	Q89U50_BRAJA	Q89u50_bradyrhizob	948	36	75.0	144	2	Q8CCL8_MOUSE	Q8ccl8_mus musculus
876	37	77.1	3122	2	P89459_HV2	P89459_human herpe	949	36	75.0	147	2	Q8BFT5_MOUSE	Q8bft5_mus muscu
877	37	77.1	3247	2	Q85533_9ALPH	Q85533_bovine herp	950	36	75.0	147	2	Q4R4K1_MACPA	Q4r4k1_macaca faec
878	37	77.1	3247	2	Q77CD4_9ALPH	Q77cd4_bovine herp	951	36	75.0	147	2	Q5JK95_ORYSA	Q5jk95_oryza sativ
879	37	77.1	3409	2	Q6SS86_CHLRE	Q6se86_chlamydomon	952	36	75.0	149	2	Q63IU4_BURPS	Q63iu4_burkholderi
880	37	77.1	3592	2	Q7W692_BORPA	Q7w692_bordetella	953	36	75.0	149	2	Q4V2Y4_BURMA	Q4v2y4_burkholderi
881	37	77.1	3634	2	Q9JP78_BORBR	Q9jp78_bordetella	954	36	75.0	151	2	Q6ZHK8_ORYSA	Q6zhk8_oryza sativ
882	37	77.1	3652	2	Q7MI61_BORBR	Q7mi61_bordetella	955	36	75.0	151	2	Q72B50_DESVH	Q72e50_desulfovibr
883	37	77.1	4027	2	Q512R0_CHLIN	Q512r0_chlamydomon	956	36	75.0	152	2	Q5VXA5_HUMAN	Q5vxa5_homo sapien
884	37	77.1	5542	2	Q7YXX2_CRYPV	Q7yxx2_cryptospori	957	36	75.0	152	2	Q9DUD0_9VIRU	Q9dud0_torque teno
885	36.5	76.0	70	2	Q7F440_ORYSA	Q7f440_oryza sativ	958	36	75.0	153	2	Q17269_BRUPA	Q17269_brugia pana
886	36.5	76.0	73	2	Q5JLV0_ORYSA	Q5jlv0_oryza sativ	959	36	75.0	153	2	Q6AGN9_LEIXX	Q6agn9_leifsonia x
887	36.5	76.0	181	2	Q75LV2_ORYSA	Q75lv2_oryza sativ	960	36	75.0	154	2	Q5Z5L8_ORYSA	Q5z5l8_oryza sativ
888	36.5	76.0	226	1	GSPB_AERHY	P45755_aeromonas h	961	36	75.0	155	2	Q6ZNT1_HUMAN	Q6znt1_homo sapien
889	36.5	76.0	651	2	Q8S587_ORYSA	Q8s5m7_oryza sativ	962	36	75.0	157	2	Q59JQ4_CANAL	Q59jq4_candida alb
890	36.5	76.0	890	1	TYR03_HUMAN	Q86418_homo sapien	963	36	75.0	157	2	Q5Z770_ORYSA	Q5z770_oryza sativ
891	36.5	76.0	890	2	Q86VR3_HUMAN	Q86vr3_homo sapien	964	36	75.0	158	2	Q6Z793_ORYSA	Q6z793_oryza sativ
892	36.5	76.0	951	2	Q59FM9_HUMAN	Q59fm9_homo sapien	965	36	75.0	158	2	Q8L490_ORYSA	Q8l490_oryza sativ
893	36.5	76.0	1085	2	Q5Z188_CHICK	Q5zie8_gallus gall	966	36	75.0	158	2	Q72KK8_THET2	Q72kk8_thermus the
894	36.5	76.0	1307	2	Q3V4J6_DROME	Q3v4j6_drosophila	967	36	75.0	158	2	Q5SKI1_THET8	Q5skil_thermus the
895	36	75.0	46	2	Q99LN4_MOUSE	Q99ln4_mus musculus	968	36	75.0	158	2	Q9DVW0_BACU	Q9dvw0_plutella xy
896	36	75.0	59	2	Q8GSH8_ORYSA	Q8gsh8_oryza sativ	969	36	75.0	159	2	Q9NCB5_HUMAN	Q9ncb5_homo sapien
897	36	75.0	60	2	Q9QVH2_9MURI	Q9qvh2_mus ep. . p	970	36	75.0	159	2	Q9NUM6_HUMAN	Q9num6_homo sapien
898	36	75.0	61	2	Q67UA3_ORYSA	Q67ua3_oryza sativ	971	36	75.0	159	2	Q4RY14_TETNG	Q4ry14_tetradodon n
899	36	75.0	64	2	Q8N507_HUMAN	Q8n5u3_homo sapien	972	36	75.0	160	2	Q9DEP6_SPAAU	Q9dep6_sparus aura
900	36	75.0	66	2	Q745U5_THET2	Q745u5_thermus the	973	36	75.0	162	2	Q55IY1_CRYNE	Q55iy1_cryptococcu
901	36	75.0	66	2	Q53WD4_THET8	Q53wd4_thermus the	974	36	75.0	162	2	Q5KCS2_CRYNE	Q5kcs2_cryptococcu
902	36	75.0	68	2	Q4WAY6_ASPFU	Q4wat6_aspergillus	975	36	75.0	164	2	Q08548_MOUSE	Q08548_mus musculus
903	36	75.0	71	2	Q4SU74_TETNG	Q4su74_tetradodon n	976	36	75.0	164	2	Q5J2U5_CRYNE	Q5j2u5_mus musculus
904	36	75.0	72	2	Q6UJ35_ORYSA	Q6uu35_oryza sativ	977	36	75.0	165	2	Q560J4_CRYNE	Q560j4_cryptococcu
905	36	75.0	73	2	Q80W36_MOUSE	Q80w36_mus musculus	978	36	75.0	165	2	Q96PC6_HUMAN	Q96pc6_homo sapien
906	36	75.0	83	2	Q96W43_OPHNO	Q96w43_ophiostoma	979	36	75.0	165	2	Q8C130_MOUSE	Q8c130_mus musculus
907	36	75.0	86	2	Q4YHU3_PLABE	Q4yhu3_plasmodium	980	36	75.0	166	2	Q6H514_ORYSA	Q6h514_oryza sativ

981 36 75.0 168 2 Q98911_CHICK
 982 36 75.0 169 2 Q9Y976_AERPE
 983 36 75.0 169 2 Q35315_mus musculus
 984 36 75.0 170 1 IR04_HCMVA
 985 36 75.0 174 2 Q9BG77_MACFA
 986 36 75.0 179 2 Q94LB7_ORISA
 987 36 75.0 182 2 Q86F22_SCHJA
 988 36 75.0 184 2 Q69K83_ORISA
 989 36 75.0 184 2 Q8BJ04_MOUSE
 990 36 75.0 187 2 Q6Z922_ORISA
 991 36 75.0 189 2 Q6F617_HUMAN
 992 36 75.0 189 2 Q8LGR6_VITVI
 993 36 75.0 192 2 Q518S5_MAGGR
 994 36 75.0 195 2 Q7QGA4_ANOGA
 995 36 75.0 195 2 Q6H4Q0_ORISA
 996 36 75.0 196 2 Q14557_HUMAN
 997 36 75.0 198 2 Q8H7G8_ARATH
 998 36 75.0 198 1 SEC65_SCHPO
 999 36 75.0 199 2 Q57XJ3_9TRYP
 1000 36 75.0 199 2 Q9N3U3_CABEL

ALIGNMENTS

RESULT 1
 Q41257_GIBZE PRELIMINARY; PRT; 1112 AA.
 AC Q41257;
 DT 13-SEP-2005 (TrEMBLrel. 31, Created)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
 DE Hypothetical protein.
 GN ORFNames=FG08701.1;
 OS Gibberella zeae PH-1.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
 OX NCBI_TaxID=229533;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=PH-1;
 RA Birren B., Nusbaum C., Abouelleil A., Allen N., Anderson S.,
 RA Arachchi H.M., Barna N., Bastien V., Bloom T., Boguslavskiy L.,
 RA Borkhachar B., Butler J., Calvo S.B., Camarata J., Chang J.,
 RA Choepel V., Collymore A., Cook A., Cooke P., Corum B., DeAtellano K.,
 RA Diaz J.S., Dodge S., Dooley K., Dorris L., Elkins T., Engels R.,
 RA Erickson J., Faro S., Ferreira P., Fitzgerald M., Gage D., Galagan J.,
 RA Gardyna S., Gnerre S., Graham L., Grand-Pierre N., Hafez N.,
 RA Hagopian D., Hagos B., Hall J., Horton L., Hulme W., Iliev I.,
 RA Jaffe D., Johnson R., Jones C., Kamal M., Kamat A., Karatas A.,
 RA Kalls C., Landers T., Levine R., Lindblad-Toh K., Liu G., Lui A.,
 RA Ma L.-J., Mabbitt R., Maclean C., MacDonald P., Major J., Manning J.,
 RA Matthews C., Mauceli E., McCarthy M., Meldrim J., Meneus L.,
 RA Mihova T., Mlenga V., Murphy T., Naylor J., Nguyen C., Nicol R.,
 RA Nielsen C.B., Norbu C., O'Connor T., O'Donnell P., O'Neill D.,
 RA Oliver J., Peterson K., Phunkhang P., Pierre N., Purcell S.,
 RA Rachupka A., Ramasamy U., Raymond C., Retta R., Rise C., Rogov P.,
 RA Roman J., Schauer S., Schuback R., Seaman S., Severy P., Smirnov S.,
 RA Smith C., Spencer B., Stange-Thomann N., Stojanovic N., Stubbs M.,
 RA Talanas J., Testaye S., Theodore J., Topham K., Travers M.,
 RA Vassiliev H., Venkataraman V.S., Viel R., Vo A., Wang S., Wilson B.,
 RA Wu X., Wyman D., Young G., Zainoun J., Zembek L., Zimmer A., Zody M.,
 RA Lander E.,
 RT "Fusarium graminearum genome sequence."
 RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
 CC -1- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 DR EMBL: AACW01000352; EAA70870.1; -; Genomic_DNA.
 KW Hypothetical protein.
 SQ SEQUENCE 1112 AA; 126572 MW; 70CE0CD407FB2C1 CRC64;

Query Match 100.0%; Score 48; DB 2; Length 1112;

Best Local Similarity 88.9%; Pred. No. 2.7e+02;
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 4 RPLPPLPXP 12
 Db 714 RPLPPLPFTP 722
 RESULT 2
 KCNA5_HUMAN
 ID KCNA5_HUMAN STANDARD; PRT; 613 AA.
 AC P22460;
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 13-SEP-2005 (Rel. 48, Last annotation update)
 DE Potassium voltage-gated channel subfamily A member 5 (Voltage-gated
 DE potassium channel subunit Kv1.5) (HK2) (HPCN1).
 GN Name=KCNA5;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
 OC Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Heart;
 RX MEDLINE=91160866; PubMed=2001794;
 RA Tamkun M.M., Knoch K.M., Walbridge J.A., Kroemer H., Roden D.M.,
 RA Glover D.M.;
 RT "Molecular cloning and characterization of two voltage-gated K+
 RT channel cDNAs from human ventricle."
 RL FASEB J. 5:331-337(1991).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=insulinoma;
 RX MEDLINE=91095456; PubMed=1986382;
 RA Philipson L.H., Hice R.E., Schaefer K., Lamendola J., Bell G.I.,
 RA Nelson D.J., Steiner D.F.;
 RT "Sequence and functional expression in Xenopus oocytes of a human
 RT insulinoma and islet potassium channel.";
 RL Proc. Natl. Acad. Sci. U.S.A. 88:53-57(1991).
 RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Heart;
 RX MEDLINE=92241872; PubMed=1349297;
 RA Curran M.E., Landes G.M., Keating M.T.;
 RT "Molecular cloning, characterization, and genomic localization of a
 RT human potassium channel gene."
 RL Genomics 12:729-737(1992).
 CC -1- FUNCTION: Mediates the voltage-dependent potassium ion
 CC permeability of excitable membranes. Assuming opened or closed
 CC conformations in response to the voltage difference across the
 CC membrane, the protein forms a potassium-selective channel through
 CC which potassium ions may pass in accordance with their
 CC electrochemical gradient. May play a role in regulating the
 CC secretion of insulin in normal pancreatic islets.
 CC -1- SUBUNIT: Heterotetramer of potassium channel proteins. Interacts
 CC via its PDZ-binding motif with DLG1. Forms a ternary complex with
 CC DLG1 and CAV3 (By similarity).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: Pancreatic islets and insulinoma.
 CC -1- DOMAIN: The amino terminus may be important in determining the
 CC rate of inactivation of the channel while the C-terminal PDZ-
 CC binding motif may play a role in modulation of channel activity
 CC and/or targeting of the channel to specific subcellular
 CC compartments (By similarity).
 CC -1- DOMAIN: The segment S4 is probably the voltage-sensor and is
 CC characterized by a series of positively charged amino acids at
 CC every third position.
 CC -1- SIMILARITY: Belongs to the potassium channel family. A (Shaker)
 CC subfamily.
 CC
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.

```
-----
CC EMBL; M60451; AAA61276.1; -; mRNA.
CC DR EMBL; M55513; AAA36422.1; -; mRNA.
CC DR EMBL; M83254; AAA60146.1; -; mRNA.
CC DR PIR; A56031; A56031.
CC DR HSP; P08510; IHO2.
CC DR SMI; P22460; 120-218.
CC DR Ensembl; ENSG00000130037; Homo sapiens.
CC DR HGNC; HGNC:6224; KCNA5.
CC DR MIM; 176267; -.
CC DR GO; GO:0008076; C: voltage-gated potassium channel complex; TAS.
CC DR GO; GO:0005281; F: delayed rectifier potassium channel activity; TAS.
CC DR GO; GO:0006813; P: potassium ion transport; TAS.
CC DR InterPro; IPR000210; BTB_POZ.
CC DR InterPro; IPR005821; Ion trans.
CC DR InterPro; IPR001622; K+channel_pore.
CC DR InterPro; IPR003091; K_channel.
CC DR InterPro; IPR003131; K_tetra.
CC DR InterPro; IPR004052; KV15channel.
CC DR InterPro; IPR003968; Kv channel.
CC DR InterPro; IPR005820; M+channel_nlg.
CC DR InterPro; IPR003972; Shaker_channel.
CC DR Pfam; PF00520; Ion_trans; 1.
CC DR Pfam; PF02214; K_tetra; 1.
CC DR PRINTS; PR00169; KCHANNEL.
CC DR PRINTS; PR01512; KV15CHANNEL.
CC DR PRINTS; PR01491; KVCHANNEL.
CC DR PRINTS; PR01496; SHAKERCHANNEL.
CC DR SMART; SM00225; BTB; 1.
KW Glycoprotein; Ion transport; Ionic channel; Lipoprotein;
KW Multigene family; Palmitate; Phosphorylation; Potassium;
KW Potassium channel; Potassium transport; Repeat; Transmembrane;
KW Transport; Voltage-gated channel.
FT TRANSMEM 251 269 Segment S1.
FT TRANSMEM 325 345 Segment S2.
FT TRANSMEM 356 377 Segment S3.
FT TRANSMEM 398 419 Segment S4.
FT TRANSMEM 434 455 Segment S5.
FT TRANSMEM 495 516 Segment S6.
FT REPEAT 61 71 1.
FT REPEAT 72 82 2.
FT REGION 61 82
FT 2 X 11 AA tandem repeat of D-[SP]-G-V-R-
FT P-L-P-P-L-P.
FT Selectivity filter (By similarity).
FT PDZ-binding (By similarity).
FT Poly-Glu.
FT Phosphoserine (by PKA) (Potential).
FT S-palmitoyl cysteine (Potential).
FT N-linked (GlcNAc...) (Potential).
FT N-linked (GlcNAc...) (Potential).
FT Missing (in Ref. 1).
FT L -> Q (in Ref. 2).
FT P -> R (in Ref. 3).
FT R -> G (in Ref. 1).
FT R -> V (in Ref. 2).
FT P -> V (in Ref. 2).
FT L -> V (in Ref. 3).
FT P -> A (in Ref. 1).
FT GPVLEKCNVRAKSNVDRSLVALCLDTSRETDL -> QL
FT PPREV (in Ref. 3).
CC SEQUENCE 613 AA; 67169 MW; 6324F704F127C5BC CRC64;
```

Query Match 97.9%; Score 47; DB 1; Length 613;
Best Local Similarity 88.9%; Pred. NO. 2e+02;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 RPLPPLPXP 12
|||||||

```
Db 65 RPLPPLPDP 73
RESULT 3
Q4VAJ1 HUMAN
ID Q4VAJ1 HUMAN PRELIMINARY; PRT; 613 AA.
AC Q4VAJ1.
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Potassium voltage-gated channel, shaker-related subfamily, member
DE 5.
GN Name=KCNA5;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RC NUCLEOTIDE SEQUENCE.
RX TISSUE=PCR rescued clones; DOI=10.1073/pnas.242603899;
RA Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.W., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,
RA Schnurch A., Schein J.B., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RL and mouse cDNA sequences.";
RN [2]
RC NUCLEOTIDE SEQUENCE.
RX TISSUE=PCR rescued clones;
RG NIH MGC Project;
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
RN [3]
RC NUCLEOTIDE SEQUENCE.
RX TISSUE=PCR rescued clones;
RG NIH MGC Project;
RL Submitted (JUL-2005) to the EMBL/GenBank/DBJ databases.
RN [4]
RC NUCLEOTIDE SEQUENCE.
RX TISSUE=PCR rescued clones;
RG NIH MGC Project;
RL Submitted (JUL-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC096358; AAH96358.1; -; mRNA.
DR EMBL; BC099665; AAH99665.1; -; mRNA.
DR EMBL; BC099666; AAH99666.1; -; mRNA.
DR EMBL; BC096357; AAH96357.1; -; mRNA.
DR InterPro; IPR000210; BTB_POZ.
DR InterPro; IPR005821; Ion trans.
DR InterPro; IPR001622; K+channel_pore.
DR InterPro; IPR003091; K_channel.
DR InterPro; IPR003131; K_tetra.
DR InterPro; IPR004052; KV15channel.
DR InterPro; IPR003968; Kv channel.
DR InterPro; IPR005820; M+channel_nlg.
DR InterPro; IPR003972; Shaker channel.
DR Pfam; PF00520; Ion_trans; 1.
DR Pfam; PF02214; K_tetra; 1.
DR PRINTS; PR00169; KCHANNEL.
DR PRINTS; PR01512; KV15CHANNEL.
DR PRINTS; PR01491; KVCHANNEL.
DR PRINTS; PR01496; SHAKERCHANNEL.
DR SMART; SM00225; BTB; 1.
KW Ion transport; Ionic channel; Transmembrane; Transport.
```

SQ SEQUENCE 613 AA; 67228 MW; A5B02B27F8396E3D CRC64;

Query Match 97.9%; Score 47; DB 2; Length 613;
Best Local Similarity 88.9%; Pred. No. 2e+02;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 RPLPPLPXP 12
|||||||
Db 65 RPLPPLPDP 73

RESULT 4
TNFL6_PIG STANDARD; PRT; 282 AA.
AC Q9BAE8; Q95M04; Q95N10;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE [Contains: Tumor necrosis factor ligand superfamily member 6 (FAS antigen ligand)
DE form; Tumor necrosis factor ligand superfamily member 6, soluble
DE form].
GN Name=TNFRSF6; Synonyms=PASL;
OS Sus scrofa (pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;
OC Sus
OX NCBI_TaxID=9923;
RN [1]
RP MEDLINE=21322533; PubMed=11429163; DOI=10.1089/107999001300177493;
RA Muneta Y., Shimoi Y., Inumaru S., Mori Y.;
RT "Molecular cloning, characterization, and expression of porcine Fas
RT ligand (CD95 ligand).";
RL J. Interferon Cytokine Res. 21:305-312(2001).
RN [2]
RP STRAIN=Guanxi bama miniature pig;
RA Zhu N., Young Y.;
RT "Molecular cloning and characterization of porcine Fas ligand cDNA.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP TISSUE=Lymphoid;
RA Tsuyuki S., Kono M., Bloom E.T.;
RT "Cloning and potential utility of porcine Fas ligand: overexpression
RT in porcine cells protects them from attack by human cytolytic cells.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
RN [4]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Landrace x Large Yorkshire white; TISSUE=Thymocyte;
RX MEDLINE=21653191; PubMed=11792426; DOI=10.1016/S0161-5890(01)00098-0;
RA Motegi-Ishiyama Y., Nakajima Y., Hoka S., Takagaki Y.;
RT "Porcine Fas-ligand gene: Genomic sequence analysis and comparison
RT with human gene.";
RL Mol. Immunol. 38:581-596(2002).
CC -1- FUNCTION: Cytokine that binds to TNFRSF6/FAS, a receptor that
CC transduces the apoptotic signal into cells. May be involved in
CC cytotoxic T cell mediated apoptosis and in T cell development.
CC TNFRSF6/FAS-mediated apoptosis may have a role in the induction of
CC peripheral tolerance, in the antigen-stimulated suicide of mature
CC T cells, or both. Binding to the decoy receptor TNFRSF6B/DCR3
CC modulates its effects (By similarity).
CC -1- SUBUNIT: Homotrimer (Probable).
CC -1- SUBCELLULAR LOCATION: Type II membrane protein and secreted (By
CC similarity).
CC -1- INDUCTION: By IL-18.
CC -1- PTM: The soluble form derives from the membrane form by
CC proteolytic processing (By similarity).
CC -1- SIMILARITY: Belongs to the tumor necrosis factor family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation

CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; AB027297; BAB40919.1; -; mRNA.
DR EMBL; AY033634; AAK56449.1; -; mRNA.
DR EMBL; AF397407; AAK84408.1; -; mRNA.
DR EMBL; AB069764; BAB64291.1; -; Genomic_DNA.
DR HSP; P01375; ATSV
DR InterPro; IPR008064; Fas_ligand.
DR InterPro; IPR006053; TNF_abc.
DR InterPro; IPR002961; TNF_C.
DR InterPro; IPR006052; TNF_family.
DR InterPro; IPR003636; TNF_subf.
DR PANTHER; PTHR15161; Fas_ligand; 1.
DR Pfam; PF00229; TNF; 1.
DR PRINTS; PRO1681; FASLIGAND.
DR PRINTS; PRO1234; TNESCROSISFCT.
DR PRINTS; PRO1237; TNFC.
DR ProDom; PD002012; TNF_subf; 1.
DR SMART; SM00207; TNF; 1.
DR PROSITE; PS00251; TNF_1; 1.
DR PROSITE; PSS0049; TNF_2; 1.
KW Apoptosis; Cytokine; Glycoprotein; Signal-anchor; Transmembrane.
FT CHAIN 1 282
FT Tumor necrosis factor ligand superfamily
FT member 6, membrane form.
FT Tumor necrosis factor ligand superfamily
FT member 6, soluble form (By similarity).
FT Signal-anchor for type II membrane
FT protein (Potential).
FT Extracellular (Potential).
FT Pro-rich.
FT Poly-Pro.
FT Cleavage (By similarity).
FT N-linked (GLNAC. .) (Potential).
FT N-linked (GLNAC. .) (Potential).
FT N-linked (GLNAC. .) (Potential).
FT N-linked (GLNAC. .) (Potential).
FT Potential.
FT F -> L (in Ref. 4).
FT T -> P (in Ref. 2).
SQ SEQUENCE 282 AA; 31756 MW; 6743DAA1145671FB CRC64;
Query Match 95.8%; Score 46; DB 1; Length 282;
Best Local Similarity 88.9%; Pred. No. 1.2e+02;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 RPLPPLPXP 12
|||||||
Db 62 RPLPPLPDP 70

RESULT 5
Q552A2 CRYNE
ID Q552A2 CRYNE PRELIMINARY; PRT; 1794 AA.
AC Q552A2;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=CNBA5860;
OS Cryptococcus neoformans var. neoformans B-3501A.
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Heterobasidiomycetes;
OC Tremellomycetidae; Tremellales; Tremellaceae; Filobasidiella.
OX NCBI_TaxID=283643;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=B-3501A;
RA Fung B., Hyman R.W., Rowley D., Miranda M., Fukushima M.,
RA Wickes B.L., Fu J., Davis R.W.;
RT "Cryptococcus neoformans serotype D sequencing.";
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an

CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC Preliminary data.
CC -1- SIMILARITY: Contains 1 PH domain.
DR EMBL; AA001000004; EAL23061.1; -; Genomic_DNA.
DR InterPro; IPR000977; DNA_ligase.
DR InterPro; IPR001849; PH.
DR InterPro; IPR000904; Sec7.
DR Pfam; PF00169; PH; 1.
DR Pfam; PF01369; Sec7; 1.
DR SMART; SM00233; PH; 1.
DR SMART; SM00222; Sec7; 1.
DR PROSITE; PS00697; DNA_LIGASE_A1; UNKNOWN_1.
DR PROSITE; PS50003; PH DOMAIN; 1.
DR PROSITE; PS50190; SEC7; 1.
KW Coiled coil; Guanine-nucleotide releasing factor;
KW Hypothetical protein.
SQ SEQUENCE 1794 AA; 136453 MW; 98811FD2F194C994 CRC64;

Query Match 95.8%; Score 46; DB 2; Length 1794;
Best Local Similarity 88.9%; Pred. No. 8.3e+02;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 RPLPPLPXP 12
Db 237 RPLPPLPPP 245

RESULT 6
Q5KNL7 CRYNE
ID Q5KNL7 CRYNE PRELIMINARY; PRT; 1811 AA.
AC Q5KNL7;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE ARF guanyl-nucleotide exchange factor, putative.
GN ORFNames=CNA06050;
OS Cryptococcus neoformans var. neoformans JEC21.
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Heterobasidiomycetes;
OC Tremellomycetidae; Tremellales; Tremellaceae; Filobasidiella.
OX NCBI_TaxID=214684;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=JEC21;
RA Loftus B.J., Fung E., Roncaglia P., Rowley D., Amedeo P., Bruno D.,
RA Vamathevan J., Miranda M., Anderson I.J., Fraser J.A., Allen J.E.,
RA Bosdet I.E., Brent M.R., Chiu R., Doering T.L., Donlin M.J.,
RA D'Souza C.A., Fox D.S., Grinberg V., Fu J., Fukushima M., Haas B.J.,
RA Huang J.C., Jambon G., Jones S.J.M., Krzywinaki M.I., Kwon-Chung J.K.,
RA Lengeler K.B., Maiti R., Marra M.A., Marra R.E., Mathewson C.A.,
RA Mitchell T.G., Perle M., Riggs F.R., Salzberg S.L., Shvartsbeyn A.,
RA Schein J.E., Shin H., Specht C.A., Suh B., Tenney A., Utterback T.,
RA Wickes B.L., Wye N.H., Kronstad J., Lodge J.K., Heitman J.,
RA Davis R.W., Fraser C.M., Hyman R.W.;
RA "The genome and transcriptome of Cryptococcus neoformans, a
RT basidiomycete fungal pathogen of humans.";
RL Science 0:0-0(2005).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=JEC21;
RA Loftus B.J., Fung E., Roncaglia P., Rowley D., Amedeo P., Bruno D.,
RA Vamathevan J., Miranda M., Anderson I.J., Fraser J.A., Allen J.E.,
RA Bosdet I.E., Brent M.R., Chiu R., Doering T.L., Donlin M.J.,
RA D'Souza C.A., Fox D.S., Grinberg V., Fu J., Fukushima M., Haas B.J.,
RA Huang J.C., Jambon G., Jones S.J.M., Koo H.L., Krzywinaki M.I.,
RA Kwon-Chung J.K., Lengeler K.B., Maiti R., Marra M.A., Marra R.E.,
RA Mathewson C.A., Mitchell T.G., Perle M., Riggs F.R., Salzberg S.L.,
RA Vamathevan J., Vamathevan J., Utterback T.,
RA Van Aken S., Fraser C.;
RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=JEC21;
RX PubMed=15553466; DOI=10.1126/science.1103773;
RA Loftus B.J., Fung E., Roncaglia P., Rowley D., Amedeo P., Bruno D.,
RA Vamathevan J., Miranda M., Anderson I.J., Fraser J.A., Allen J.E.,
RA Bosdet I.E., Brent M.R., Chiu R., Doering T.L., Donlin M.J.,
RA D'Souza C.A., Fox D.S., Grinberg V., Fu J., Fukushima M., Haas B.J.,
RA Huang J.C., Jambon G., Jones S.J.M., Koo H.L., Krzywinaki M.I.,
RA Kwon-Chung J.K., Lengeler K.B., Maiti R., Marra M.A., Marra R.E.,
RA Mathewson C.A., Mitchell T.G., Perle M., Riggs F.R., Salzberg S.L.,
RA Vamathevan J., Vamathevan J., Utterback T.,
RA Van Aken S., Fraser C.;
RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.
RN [3]

RA Schein J.E., Shvartsbeyn A., Shin H., Shumway M., Specht C.A.,
RA Suh B.B., Tenney A., Utterback T.R., Wickes B.L., Wortman J.R.,
RA Wye N.H., Kronstad J.W., Lodge J.K., Heitman J., Davis R.W.,
RA "The genome of the basidiomycetous yeast and human pathogen
RT Cryptococcus neoformans.";
RL Science 307:1321-1324(2005).
CC -1- SIMILARITY: Contains 1 PH domain.
DR EMBL; AE017341; ANM41140.1; -; Genomic_DNA.
DR InterPro; IPR000977; DNA_ligase.
DR InterPro; IPR001849; PH.
DR InterPro; IPR011993; PH type.
DR InterPro; IPR000904; Sec7.
DR Pfam; PF00169; PH; 1.
DR Pfam; PF01369; Sec7; 1.
DR SMART; SM00233; PH; 1.
DR SMART; SM00222; Sec7; 1.
DR PROSITE; PS00697; DNA_LIGASE_A1; UNKNOWN_1.
DR PROSITE; PS50003; PH DOMAIN; 1.
DR PROSITE; PS50190; SEC7; 1.
KW Coiled coil; Complete proteome; Guanine-nucleotide releasing factor.
SQ SEQUENCE 1811 AA; 198207 MW; 33888A8C9C24019 CRC64;

Query Match 95.8%; Score 46; DB 2; Length 1811;
Best Local Similarity 88.9%; Pred. No. 8.4e+02;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 RPLPPLPXP 12
Db 237 RPLPPLPPP 245

RESULT 7
FDSCP_HUMAN
ID FDSCP_HUMAN STANDARD; PRT; 85 AA.
AC Q8NFU4;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Follicular dendritic cell secreted peptide precursor (FDC-SP) (FDC
DE secreted protein).
GN Name=FDCSP; Synonyms=C4orf7; ORFNames=UNQ733/PRO1419;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE, AND TISSUE SPECIFICITY.
RC TISSUE=Tonsil;
RX MEDLINE=22181501; PubMed=12193705;
RA Marshall A.J., Du Q., Draves K.E., Shikishima Y., Hayglass K.T.,
RA Clark E.A.;
RA "FDC-SP, a novel secreted protein expressed by follicular dendritic
RT cells.";
RL J. Immunol. 169:2381-2389(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RX MEDLINE=22887296; PubMed=12975309; DOI=10.1101/gr.1293003;
RA Clark H.F., Gurney A.L., Abaya E., Baker K., Baldwin D.T., Brush J.,
RA Chen J., Chow B., Chui C., Crowley C., Currell B., Deuel B., Dowd P.,
RA Eaton D., Foster J.S., Grimaldi C., Gu Q., Haas P.E., Heidens S.,
RA Huang A., Kim H.S., Klimowski L., Jin Y., Johnson S., Lee J.,
RA Lewis L., Liao D., Mark M.R., Robbie E., Sanchez C., Schoenfeld J.,
RA Seahagiri S., Simmons L., Singh J., Smith V., Stinson J., Vagts A.,
RA Vandlen R.L., Watanabe C., Wleand D., Woods K., Xie M.-H.,
RA Yansura D.G., Yi S., Yu G., Yuan J., Zhang M., Zhang Z., Goddard A.D.,
RA Wood W.I., Godowski P.J., Gray A.M.;
RA "The secreted protein discovery initiative (SPDI), a large-scale
RT effort to identify novel human secreted and transmembrane proteins: a
RL bioinformatics assessment.";
RL Genome Res. 13:2265-2270(2003).
RN [3]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC TISSUE=Thyroid;
RX MEDLINE=23289257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Fellings E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haig F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bobak S.A., McWen P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Haley S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Vallalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.B., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences".
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
CC -!- FUNCTION: Can bind to the surface of B lymphoma cells, but not T
CC lymphoma cells, consistent with a function as a secreted mediator
CC acting upon B cells.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Abundantly expressed in tonsil, lymph node,
CC and thymus; strong expression in prostate; lower expression in
CC thyroid, stomach, and colon.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL: AF435080; AAN01116.1; -; mRNA.
DR EMBL: AY358958; AA089317.1; -; mRNA.
DR EMBL: BC062213; AA062213.1; -; mRNA.
DR Ensembl: ENSG00000181617; Homo sapiens.
DR HGNC: HGNC:19215; C4orf7.
DR MIM: 607241; -.
KW Signal.
FT SIGNAL 1 17 Potential.
FT CHAIN 18 85 Follicular dendritic cell secreted
FT COMPBIAS 45 82 Peptide.
FT COMPBIAS 85 AA; 9700 MW; 3FEC98BDE883AB07 CRC64;
SQ SEQUENCE 85 AA; 9700 MW; 3FEC98BDE883AB07 CRC64;
Query Match 93.8%; Score 45; DB 1; Length 85;
Best Local Similarity 77.8%; Pred. No. 46;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Qy 4 RLPPLPXP 12
Db 51 RLPPLPFP 59
RESULT 8
Q540F3 HUMAN
ID Q540F3 HUMAN PRELIMINARY; PRT; 85 AA.
AC Q540F3;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE NPC-associated peptide.
GN Name=NAP1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;

RP NUCLEOTIDE SEQUENCE.
RA Li F., Jiang W.H., Yao K.T.;
RT "Cloning and functional research of a novel human NPC-associated gene
RT similar to FDC-SP".;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY190326; AAO41714.1; -; mRNA.
SQ SEQUENCE 85 AA; 9700 MW; 3FEC98BDE883AB07 CRC64;
Query Match 93.8%; Score 45; DB 2; Length 85;
Best Local Similarity 77.8%; Pred. No. 46;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Qy 4 RLPPLPXP 12
Db 51 RLPPLPFP 59
RESULT 9
Q91NH0 9PARA
ID Q91NH0 9PARA PRELIMINARY; PRT; 519 AA.
AC Q91NH0;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Nucleocapsid protein.
OS Tioman virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Rubulavirus.
OX NCBI_TaxID=162013;
Qy 4 RLPPLPXP 12
Db 51 RLPPLPFP 59
RESULT 10
Q6ZS83 HUMAN
ID Q6ZS83 HUMAN PRELIMINARY; PRT; 147 AA.
AC Q6ZS83;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)

DE Hypothetical protein FLJ45743.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Kidney;
RA Oshima A., Takahashi-Fujii A., Tanase T., Imose N., Takeuchi K.,
RA Arita M., Musashino K., Yuuki H., Hara H., Sugiyama T., Irie R.,
RA Otsuki T., Sato H., Wakamatsu A., Iehii S., Yamamoto J., Isono Y.,
RA Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Wagatsuma M.,
RA Murakawa K., Kanehori K., Sugiyama A., Kawakami B., Suzuki Y.,
RA Sugano S., Nagahori K., Maeno Y., Nagai K., Isogai T.,
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK127645; BAC97071.1; -; mRNA.
SQ SEQUENCE 147 AA; 15916 MW; 0D82561EEF4F2A0A CRC64;

Query Match 91.7%; Score 44; DB 2; Length 147;
Best Local Similarity 77.8%; Pred. No. 1.1e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 4 RPLPPLXP 12
Db 82 RPLPPLCP 90
||:|||||

RESULT 11
Q54M51 D1CIDI
ID Q54M51 D1CIDI PRELIMINARY; PRT; 252 AA.
AC Q54M51;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=DD80218797;
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=AX4;
RA Eichinger L., Pachebat J.A., Gloeckner G., Rajandream M.-A.,
RA Xuegang R., Berriman M., Song J., Olsen R., Szafranski K., Xu Q.,
RA Tunggai B., Kummerfeld S., Madera M., Konfortov B.A., Rivero F.,
RA Bankier A.T., Lehmann R., Hamlin N., Davies R., Gaudet P., Fey P.,
RA Pilcher K., Chen G., Saunders D., Sodergren E., Davis P.,
RA Kerhornou A., Nie X., Hall N., Anjard C., Hemphill L., Bason N.,
RA Farbrother P., Desany B., Just E., Morio T., Rost R., Churcher C.,
RA Cooper J., Haydock S., van Driessche N., Cronin A., Goodhead I.,
RA Muzny D., Mourier T., Pain A., Lu M., Harper D., Lindsay R.,
RA Hauser H., James K., Quiles M., Mohan M.B., Saito T., Buchrieser C.,
RA Wardrop A., Felder M., Thangavelu M., Johnson D., Knights A.,
RA Louissegh H., Mungall K., Oliver K., Price C., Quail M.A.,
RA Urushihara H., Hernandez J., Rabinowitsch E., Steffen D., Sanders M.,
RA Ma J., Kohara Y., Sharp S., Simmonds M., Spiegler S., Tivey A.,
RA Shugano S., White B., Walker D., Woodward J., Winckler T., Tanaka Y.,
RA Shaulsky G., Schleicher M., Weinstein G., Rosenthal A., Cox E.C.,
RA Chisholm R.L., Gibbs R., Loomis W.F., Platzer M., Kay R.R.,
RA Williams J., Dear P.H., Noegel A.A., Barrell B., Kuspa A.;
RL "The genome of the social amoeba Dictyostelium discoideum";
RL Nature 0:0-0(2005).

CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC EMBL; AAF01000134; EAL64339.1; -; Genomic_DNA.
DR Hypothetical protein.
KW Hypothetical protein.
SQ SEQUENCE 252 AA; 29168 MW; A4C13A5AF49B673E CRC64;

Query Match 91.7%; Score 44; DB 2; Length 252;
Best Local Similarity 77.8%; Pred. No. 2e+02;

Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Qy 4 RPLPPLXP 12
Db 186 KPLPPLPOP 194
:|||||

RESULT 12
Q5M9F4 RAT
ID Q5M9F4 RAT PRELIMINARY; PRT; 355 AA.
AC Q5M9F4;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE RGD1306209 predicted protein (Fragment).
DE Name=RGD1306209_predicted;
GN Rattus norvegicus (Rat).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Testis;
MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

Qy 4 RPLPPLXP 12
Db 309 QPLPPLPAP 317
:|||||

Query Match 91.7%; Score 44; DB 2; Length 355;
Best Local Similarity 77.8%; Pred. No. 2.8e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Qy 4 RPLPPLXP 12
Db 309 QPLPPLPAP 317
:|||||

RESULT 13
Q5M9F4 RAT
ID Q5M9F4 RAT PRELIMINARY; PRT; 355 AA.
AC Q5M9F4;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE RGD1306209 predicted protein (Fragment).
DE Name=RGD1306209_predicted;
GN Rattus norvegicus (Rat).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Testis;
MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

Qy 4 RPLPPLXP 12
Db 309 QPLPPLPAP 317
:|||||

Query Match 91.7%; Score 44; DB 2; Length 355;
Best Local Similarity 77.8%; Pred. No. 2.8e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Qy 4 RPLPPLXP 12
Db 309 QPLPPLPAP 317
:|||||

RESULT 13

Q43516 LYCES
ID Q43516 LYCES PRELIMINARY; PRT; 162 AA.
AC Q43516;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Unknown protein (Fragment).
GN Name=unknown;
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.
OX NCBI_TaxID=4081;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Leaf;
RX MEDLINE=97238471; PubMed=9132056; DOI=10.1023/A:1005746931643;
RA John I., Hackett R., Cooper W., Drake R., Farrell A., Grierson D.,
RT "Cloning and characterization of tomato leaf senescence-related
cDNAs";
RL Plant Mol. Biol. 33:641-651(1997).
DR EMBL: Z75519; CAA9755.1; -; mRNA.
DR F1R; T07173; T07173.
FT NON TER
SQ SEQUENCE 162 AA; 17739 MW; 8C5F10D7DB2FA300 CRC64;

Query Match 89.6%; Score 43; DB 2; Length 162;
Best Local Similarity 87.5%; Pred. No. 1.7e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 PLPPLXP 12
| | | | |
Db 155 PLPPLPSP 162

RESULT 14
Q6KA56 ORYSA
ID Q6KA56 ORYSA PRELIMINARY; PRT; 162 AA.
AC Q6KA56;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein OJ1115_D03.31.
GN Name=OJ1115_D03.31;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsumoto T., Yamamoto K.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AP004001; BAD21501.1; -; Genomic_DNA.
DR Gramineae; Q6KA56; -.
KW Hypothetical protein.
SQ SEQUENCE 162 AA; 17550 MW; 7BCB70E2D249220 CRC64;

Query Match 89.6%; Score 43; DB 2; Length 162;
Best Local Similarity 87.5%; Pred. No. 1.7e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 PLPPLXP 12
| | | | |
Db 84 PLPPLPSP 91

RESULT 15
Q8BX0 MOUSE
ID Q8BX0 MOUSE PRELIMINARY; PRT; 165 AA.
AC Q8BX0;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Mus musculus 7 days neonate cerebellum cDNA, RIKEN full-length
DE enriched library, clone:A730035117 product:hypothetical Alanine-rich
DE region containing protein, full insert sequence. (Fragment).
GN Name=Prdm8;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadoya K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Giesi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schram L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seyta T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection";
RL Nature 409:685-690(2001).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RA The PANTOM Consortium.
RT The RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
60,770 full-length cDNAs";
RL Nature 420:563-573(2002).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes";
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nageoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuunai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multicapillary sequencer";
RL Genome Res. 10:1757-1771(2000).
RN [6]


```

RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Haneigaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayateu N., Hiramoto K., Hirooka T., Hirozane T.,
RA Kori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Kato H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toy T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR ENBL; AK042889; BAC1394.1; -; mRNA.
DR MGI; MGI:1924880; Prdm8.
KW Hypothetical protein.
FT NON TPR 1
SQ SEQUENCE 165 AA; 16717 MW; 8298A509CA439E5F CRC64;

Query Match 89.6%; Score 43; DB 2; Length 165;
Best Local Similarity 87.5%; Pred. No. 1.7e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 PLPLPXP 12
DB 113 PLPLPXP 120

RESULT 16
Q8H5L8 ORYSA
ID Q8H5L8_ORYSA PRELIMINARY; PRT; 170 AA.
AC Q8H5L8;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein OJ1060_D03.115.
GN Name=OJ1060_D03.115;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 7, BAC
RT clone:OJ1060_D03."
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR ENBL; AP003803; BAC22270.1; -; Genomic_DNA.
DR Gramene; Q8H5L8; -
DR GO; GO:0008289; F:lipid binding; IEA.
DR GO; GO:0006869; P:lipid transport; IEA.
DR InterPro; IPR003612; RAI.
DR InterPro; IPR000528; Plant_LTP.
DR Pfam; PF00234; Tryp_alpha_aml1; 1.
DR PRINTS; PR00382; LIPIDTRNSFER.
DR SMART; SM00499; AAI; 1.
KW Hypothetical protein.
SQ SEQUENCE 170 AA; 17765 MW; 6E78B8CB2429A88 CRC64;

Query Match 89.6%; Score 43; DB 2; Length 170;
Best Local Similarity 87.5%; Pred. No. 1.8e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 PLPLPXP 12
DB 49 PLPLPXP 56

RESULT 17
CX11_HUMAN
ID CX11_HUMAN STANDARD; PRT; 209 AA.

```

```

AC O15255;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE CAAX box protein 1 (Cerebral protein 5).
GN Name=CX11; ORFNames=hucep-5;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Brain;
RA Frattini A., Faranda S., Zucchi I., Verzoni P.;
RX MEDLINE=98066781; PubMed=9403077; DOI=10.1006/geno.1997.5006;
RT "A low-copy repeat in Xq26 represents a novel putatively prenylated
RT protein gene (CX11) and its pseudogenes (DXS9914, DXS9915, and
RT DXS9916).";
RL Genomics 46:167-169(1997).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC TISSUE=Brain;
RA Yu W., Sarginson J., Gibbs R.A.;
RN Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Brain;
RA Yoshimoto M., Yazaki M., Takayama K., Matsumoto K.;
RT "Biological functions of a novel human gene, hucep-5, which is
RT specifically expressed in the central nervous system.";
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RN [4]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX PubMed=15772651; DOI=10.1038/nature03440;
RA Ross M.T., Grafham D.V., Coffey A.J., Scherer S., McLay K., Muzny D.,
RA Platzer M., Howell G.R., Burrows C., Bird C.P., Frankish A.,
RA Lovell F.L., Howe K.L., Ashurst J.L., Fulton R.S., Soudbrak R., Wen G.,
RA Jones M.C., Hurler M.E., Andrews T.D., Scott C.E., Searle S., Chen R.,
RA Ramser J., Whitaker A., Deadman R., Carter N.P., Hunt S.E., Chen R.,
RA Cree A., Gunaratne P., Havlak P., Hodgson A., Metzker M.L.,
RA Richards S., Scott G., Steffen D., Sodergren E., Wheeler D.A.,
RA Worley K.C., Ainscough R., Ambrose K.D., Ansari-Lari M.A., Arachya S.,
RA Ashwell R.I., Babbage A.K., Bagguley C.L., Ballabio A., Banerjee R.,
RA Barker G.E., Barlow K.F., Barrett I.P., Bates K.N., Beare D.M.,
RA Beasley H., Beasley O., Beck A., Bethel G., Blechschmidt K., Brady N.,
RA Bray-Allen S., Bridgeman A.M., Brown A.J., Brown M.J., Bonnin D.,
RA Burford E.A., Buhay C., Burch P., Burford D., Burgess J., Burrill W.,
RA Burton J., Bye J.M., Carder C., Carrel L., Chako J., Chapman J.C.,
RA Chavez D., Chen E., Chen G., Chen Y., Chen Z., Chinault C.,
RA Ciccodicola A., Clark S.Y., Clarke G., Clee C.M., Clegg S.,
RA Clerc-Blankenburg K., Clifford K., Cobley V., Cole C.G., Conquer J.S.,
RA Corby N., Connor R.E., David R., Davies J., Davis C., Davis J.,
RA Delgado O., Dehazo D., Dhani P., Ding Y., Dinh H., Dodsworth S.,
RA Edes T., Ellwood M., Emery-Cohen A., Dunn M., Durbin K.J., Dutta I.,
RA Faulkner L., Francis P., Frankland J., Fraser A.E., Gallego P.,
RA Gilbert J., Gill R., Gloeckner G., Gregory S.G., Gribble S.,
RA Griffiths C., Grocock R., Gu Y., Gwilliam R., Hamilton C., Hart E.A.,
RA Hawes A., Heath P.D., Heilmann K., Hennig S., Hernandez J.,
RA Hinzmann B., Ho S., Hoffs M., Howden P.J., Huckle E.J., Hume J.,
RA Hunt P.J., Hunt A.R., Isherwood J., Jacob L., Johnson D., Jones S.,
RA de Jong P.J., Joseph S., Keenan S., Kelly S., Kershaw J.K., Khan Z.,
RA Kioschis P., Klages S., Knights A.J., Kosiura A., Kovar-Smith C.,
RA Laird G.K., Langford C., Lawlor S., Leversha M., Lewis L., Liu W.,
RA Lloyd C., Lloyd D.M., Loulseged H., Loveland J.E., Lovell J.D.,
RA Lozado R., Lu J., Lyne R., Ma J., Maeshwari M., Matthews L.H.,
RA McDowall J., McLaren S., McMurray A., Meidl P., Meitinger T.,
RA Milne S., Miner G., Mistry S.L., Morgan M., Morris S., Mueller I.,
RA Mullikin J.C., Nguyen N., Nordsiek G., Nyakatura G., O'dell C.N.,
RA Okwuonu G., Palmer S., Pandian R., Parker D., Parrish J.,
RA Pasternak S., Patel D., Pearce A.V., Pearson D.M., Pelan S.E.,
RA Perez L., Porter K.M., Ramsey Y., Reichwald K., Rhodes S.,

```

RA Ridler K.A., Schlessinger D., Schueler M.G., Sehra H.K.,
RA Shaw-Smith C., Shen H., Sheridan E.M., Shownkeen R., Skuce C.D.,
RA Smith M.L., Sothman E.C., Steingrubber H.E., Steward C.A., Storey R.,
RA Swann R.M., Swarbrick D., Tabor P.E., Taudien S., Taylor T.,
RA Teague B., Thomas K., Thorpe A., Timms K., Tracey A., Trevanion S.,
RA Tromans A.C., d'Uro M., Verduzco D., Villasana D., Waldron L.,
RA Wall M., Wang Q., Warren J., Watry G.L., Wei X., West A.,
RA Whitehead S.L., Whiteley M.N., Wilkinson J.E., Willey D.L.,
RA Williams G., Williams L., Williamson A., Williamson J., Wilming L.,
RA Woodmansley R.L., Wray P.W., Yen J., Zhang J., Zhou H., Zoghbi H.,
RA Zorilla S., Buck D., Reinhardt R., Poustka A., Rosenthal A.,
RA Lehrach H., Meindl A., Minx P.J., Hillier L.W., Willard H.F.,
RA Wilson R.K., Waterston R.H., Rice C.M., Vaudin M., Coulson A.,
RA Gibbs R.A., Beck S., Rogers J., Bentley D.R.,
RT "The DNA sequence of the human X chromosome."
RL Nature 434:325-337(2005).
[5]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC TISSUE=Muscle;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Sapletenko M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uedini T.B., Tohiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Vallalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettaman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalley D.E.,
RA Schnarch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- SUBCELLULAR LOCATION: Attached to the plasma membrane (Potential).
CC -1- TISSUE SPECIFICITY: Ubiquitous.
CC -1- CAUTION: Ref.5 sequences differ from that shown due to wrong
choices of frame.

CC This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.

DR EMBL; Y13374; CA473804.1; -; mRNA.
DR EMBL; AF038168; AAB97361.1; -; mRNA.
DR EMBL; AF052096; AAC28618.1; -; mRNA.
DR EMBL; D88756; BAB46920.1; -; mRNA.
DR EMBL; AL136169; CAB87958.1; -; Genomic DNA.
DR EMBL; BC002385; AAH02385.2; ALT SEQ; mRNA.
DR EMBL; BC002410; AAH02410.2; ALT SEQ; mRNA.
DR Ensembl; ENSG00000134590; Homo sapiens.
DR HGNC; HGNC:2569; CXX1.
DR H-InvDB; HIX0017070; -.
DR MIM; 300213; -.
KM Lipoprotein; Membrane; Prenylation.
FT Lipid 206 206 S-farnesyl cysteine (Potential).
SQ SEQUENCE 209 AA; 22278 MW; 976E0D2AE93023A8 CRC64;

Query Match 89.6%; Score 43; DB 1; Length 209;
Best Local Similarity 87.5%; Pred. NO. 2.2e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
|||||
QY 5 PLPPLPXP 12

Db 119 PLPPLPSP 126

RESULT 18
Q6TBF1_HUMAN
ID Q6TBF1_HUMAN PRELIMINARY; PRT; 209 AA.
AC Q6TBF1;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE CXX1 protein.
GN Name=CXX1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Ebert L., Schick M., Neubert P., Schatten R., Henze S., Korn B.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR456853; CAG33134.1; -; mRNA.
SQ SEQUENCE 209 AA; 22277 MW; 976E0D2AE93023A8 CRC64;

Query Match 89.6%; Score 43; DB 2; Length 209;
Best Local Similarity 87.5%; Pred. NO. 2.2e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
|||||
QY 5 PLPPLPXP 12
Db 119 PLPPLPSP 126

RESULT 19
Q51CQ5_ENTHI
ID Q51CQ5_ENTHI PRELIMINARY; PRT; 238 AA.
AC Q51CQ5;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=19.t00018;
OS Entamoeba histolytica HM-1:IMSS.
OC Eukaryota; Entamoebidae; Entamoeba.
OX NCBI_TaxID=294381;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=HM-1:IMSS;
RX PubMed=15729342; DOI=10.1038/nature03291;
RA Loftus B., Anderson I., Davies R., Almark U.C., Samuleson J.,
RA Amedeo P., Roncaglia P., Beriman M., Hitt R.P., Mann B.J., Nozaki T.,
RA Suh B., Pop M., Duchene M., Ackers J., Tammich E., Leippe M.,
RA Hofer M., Bruchhaus I., Willhoft U., Bhattacharya A.,
RA Chillingworth T., Churcher C., Hance Z., Harris B., Harris D.,
RA Jagels K., Moule S., Mungall K., Ormond D., Squares R., Whitehead S.,
RA Quail M.A., Rabinowitsch E., Norbertczak H., Price C., Wang Z.,
RA Guillen N., Gilchrist C., Stroup S.E., Bhattacharya S., Lohia A.,
RA Foster P.G., Sicheritz-Ponten T., Weber C., Singh U., Mukherjee C.,
RA El-Sayed N.M., Petri W.A., Clark C.G., Embley T.M., Barrell B.,
RA Fraser C.M., Hall N.;
RT "The genome of the protist parasite Entamoeba histolytica";
RL Nature 433:865-868(2005).
CC -1- CAUTION: The sequence shown here is derived from an
EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
preliminary data.
CC EMBL; AAP01000091; EAL50578.1; -; Genomic DNA.
DR EMBL; AAP01000091; EAL50578.1; -; Genomic DNA.
KW Hypothetical protein.
SQ SEQUENCE 238 AA; 27384 MW; 3EAL3C9FC5259A8B CRC64;

Query Match 89.6%; Score 43; DB 2; Length 238;
Best Local Similarity 77.8%; Pred. NO. 2.5e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
|||||

```

QY      4 RPLPLPXP 12
Db      124 RPLPPTTP 132

RESULT 20
Q4QHF1 LEIMA
ID Q4QHF1 LEIMA PRELIMINARY; PRT; 243 AA.
AC Q4QHF1
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=lmjF10.0640;
OS Leishmania major.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Friedlin;
RA Peacock C.S., Murphy L., Ivens A.C., Berriman M., Blackwell J.,
RA Smith D., Collins M., Foster N., Harris D., Oliver K., O'Neill S.,
RA Saunders D., Seeger K., Warren T., Rajandream M., and Barrell B.G.;
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; CT005249; CAJ02703.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 243 AA; 25786 MW; 9CACE2038EDF767E CRC64;

Query Match 89.6%; Score 43; DB 2; Length 243;
Best Local Similarity 87.5%; Pred. No. 2.6e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      5 PLPPLPXP 12
Db      151 PLPPLPSP 158

RESULT 21
UDU3_ARATH
ID UDU3_ARATH STANDARD; PRT; 287 AA.
AC Q9FHD3;
DT 29-MAR-2004 (Rel. 43, Created)
DT 29-MAR-2004 (Rel. 43, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Hypothetical DUF26 domain containing protein 3 precursor.
GN OrderedLocuNames=At5g41300; ORFNames=KIO13.10;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=cv. Columbia;
RX MEDLINE=20181125; PubMed=10718197;
RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Kotani H.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence
RT features of the regions of 3,076,755 bp covered by sixty P1 and TAC
RT clones."
RL DNA Res. 7:31-63(2000).
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor
CC (Potential).
CC -1- SIMILARITY: Contains 2 DUF26 domains.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; AB019225; BAB11106.1; -; Genomic_DNA.
DR GeneFarm; 704; 92.

```

```

DR InterPro; IPR002902; DUF26.
DR Pfam; PF01657; DUF26; 2.
KW Glycoprotein; GPI-anchor; Hypothetical protein; Lipoprotein; Membrane;
KW Repeat; Signal.
FT SIGNAL 1 26 Potential.
FT CHAIN 27 269 Hypothetical DUF26 domain containing
FT protein 3.
FT PROPEP 270 287 Removed in mature form (Potential).
FT DOMAIN 82 136 DUF26 1.
FT DOMAIN 198 251 DUF26 2.
FT LIPID 269 269 GPI-anchor amidated glycine (Potential).
FT CARBOHYD 43 43 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 47 47 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 63 63 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 72 72 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 93 93 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 103 103 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 111 111 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 212 212 N-linked (GlcNAc...) (Potential).
SQ SEQUENCE 287 AA; 32475 MW; B429DCDF9BAB1B8F CRC64;

Query Match 89.6%; Score 43; DB 1; Length 287;
Best Local Similarity 87.5%; Pred. No. 3.1e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      5 PLPPLPXP 12
Db      251 PLPPLPSP 258

RESULT 22
Q9FBL5_9ACTO
ID Q9FBL5_9ACTO PRELIMINARY; PRT; 462 AA.
AC Q9FBL5;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Peptide synthetase NRPS.
GN Name=Dmlii;
OS Streptomyces verticillus.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=29309;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC15003;
RX MEDLINE=20389599; PubMed=10930733;
RA Du L., Chen M., Sanchez C., Shen B.;
RT "An oxidation domain in the BmIII non-ribosomal peptide synthetase
RT probably catalyzing thiazole formation in the biosynthesis of the
RT anti-tumor drug bleomycin in Streptomyces verticillus ATCC15003."
RL FEMS Microbiol. Lett. 189:171-175(2000).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC15003;
RX MEDLINE=20500448; PubMed=11048953; DOI=10.1016/S1074-5521(00)00011-9;
RA Du L., Sanchez C., Chen M., Edwards D.J., Shen B.;
RT "The biosynthetic gene cluster for the antitumor drug bleomycin from
RT Streptomyces verticillus ATCC15003 supporting functional interactions
RT between nonribosomal peptide synthetases and a polyketide synthase."
RL Chem. Biol. 7:623-642(2000).
DR EMBL; AF210249; AAG02367.1; -; Genomic_DNA.
DR InterPro; IPR001242; Condensatn.
DR Pfam; PF00668; Condensatn; 1.
SQ SEQUENCE 462 AA; 49808 MW; 11593C8433F96D00 CRC64;

Query Match 89.6%; Score 43; DB 2; Length 462;
Best Local Similarity 87.5%; Pred. No. 5.1e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      5 PLPPLPXP 12
Db      447 PLPPLPAP 454

```

Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

RESULT 23
Q9KW38_9RICK
ID Q9KW38_9RICK PRELIMINARY; PRT; 486 AA.
AC Q9KW38;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Component of type IV secretion system.
GN Name=VirB10;
OS Wolbachia sp. wKueVO.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Rickettsiaceae; Wolbachiae; Wolbachia.
OX NCBI_TaxID=130485;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MKueVO;
RA Masui S., Sasaki T., Ishikawa H.;
RT "Genes for the Type IV Secretion System in an Intracellular Symbiont,
RT Wolbachia, a Causative Agent of Various Sexual Alterations in
RT Arthropods.";
RL J. Bacteriol. 182:6529-6531 (2000).
DR EMBL; AB045235; BAA97441.1; -; Genomic DNA.
DR GO; GO:0009291; P:unidirectional conjugation; IEA.
DR InterPro; IPR005498; Trb1.
DR Pfam; PF03743; Trb1; 1.
SQ SEQUENCE 486 AA; 53026 MW; 590D0DE79240DEBF CRC64;

Query Match 89.6%; Score 43; DB 2; Length 486;
Best Local Similarity 87.5%; Pred. No. 5.3e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 PLPPLPXP 12
Db 95 PLPPLPXP 102

RESULT 24
Q73IY6_WOLPM
ID Q73IY6_WOLPM PRELIMINARY; PRT; 486 AA.
AC Q73IY6;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Type IV secretion system protein VirB10.
GN OrderedLocNames=WD0006;
OS Wolbachia pipientis wMel.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Rickettsiaceae; Wolbachiae; Wolbachia.
OX NCBI_TaxID=66077;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX PubMed=15024419; DOI=10.1371/journal.pbio.0020069;
RA Wu M., Sun L.V., Vamathevan J.J., Riegler M., DeBoy R.T.,
RA Brownlie J.C., McGraw E.A., Martin W., Esser C., Ahmadinejad N.,
RA Wiegand C., Madupu R., Beanan M.J., Brinkac L.M., Daugherty S.C.,
RA Durkin A.S., Kolonay J.F., Nelson W.C., Mohamoud Y., Lee P.,
RA Barry K.J., Young M.B., Uterback T.R., Meidman J.F., Nierman W.C.,
RA Paulsen I.T., Nelson K.E., Tettelin H., O'Neill S.L., Eisen J.A.;
RT "Phylogenomics of the reproductive parasite Wolbachia pipientis wMel:
RT a streamlined genome overrun by mobile genetic elements.";
RL PLoS Biol. 2:327-341 (2004).
DR EMBL; AE017256; AAS13774.1; -; Genomic DNA.
DR TIGR; WD0006; -.
DR GO; GO:0009291; P:unidirectional conjugation; IEA.
DR InterPro; IPR005498; Trb1.
DR Pfam; PF03743; Trb1; 1.
KW Complete proteome.
SQ SEQUENCE 486 AA; 53012 MW; 5913FC000F69B26F CRC64;

Query Match 89.6%; Score 43; DB 2; Length 486;
Best Local Similarity 87.5%; Pred. No. 5.3e+02;

Qy 5 PLPPLPXP 12
Db 95 PLPPLPXP 102

RESULT 25
Q63IB6_BURPS
ID Q63IB6_BURPS PRELIMINARY; PRT; 487 AA.
AC Q63IB6;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Putative MFS family transporter.
GN OrderedLocNames=APS2155;
OS Burkholderia pseudomallei (Pseudomonas pseudomallei).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Burkholderia; pseudomallei group.
OX NCBI_TaxID=28450;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=X96243;
RX PubMed=15377794; DOI=10.1073/pnas.0403302101;
RA Holden M.T.G., Titball R.W., Peacock S.J., Cerdeno-Tarraga A.-M.,
RA Atkins T., Crossman L.C., Pitt T., Churcher C., Mungall K.L.,
RA Bentley S.D., Sebahia M., Thomson N.R., Bason N., Beacham I.R.,
RA Brooks K., Brown K.A., Brown N.F., Challis G.L., Cherevach I.,
RA Chillingworth T., Cronin A., Crosssett B., Davis P., Deshazer D.,
RA Feltwell T., Fraser A., Hance Z., Hauser H., Holroyd S., Jageis K.,
RA Keith K.E., Maddison M., Moule S., Price C., Quail M.A.,
RA Rabinovitsch E., Rutherford K., Sanders M., Simmonds M.,
RA Songaivilai S., Stevens K., Tumapa S., Vesaratchaveat M.,
RA Whitehead S., Yeate C., Barrell B.G., Oyston P.C.F., Parkhill J.;
RT "Genomic plasticity of the causative agent of melioidosis,
RT Burkholderia pseudomallei.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:14240-14245 (2004).
DR EMBL; BX571966; CAH39639.1; -; Genomic DNA.
DR GO; GO:0019866; C:inner membrane; IEA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR011701; MFS 1.
DR InterPro; IPR005829; Sug_transporter.
DR Pfam; PF07690; MFS 1; 1.
DR PROSITE; PS00850; MFS; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 487 AA; 51782 MW; E2F6F509629CDB0C CRC64;

Query Match 89.6%; Score 43; DB 2; Length 487;
Best Local Similarity 77.8%; Pred. No. 5.3e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 4 RPLPPLPXP 12
Db 244 RPLPPLPXP 252

RESULT 26
Q4P068_USTMA
ID Q4P068_USTMA PRELIMINARY; PRT; 488 AA.
AC Q4P068;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=UM06495.1;
OS Ustilago maydis 521.
OC Eukaryota; Fungi; Basidiomycota; Ustilaginomycetes;
OC Ustilaginomycetidae; Ustilaginales; Ustilaginaceae; Ustilago.
OX NCBI_TaxID=237631;

```

RN NUCLEOTIDE SEQUENCE.
RP STRAIN=521;
RA Alt-zahra M., Allen T., Allen T., An P., Anderson M., Anderson S.,
RA Arachi H., Armstrong J., Bachantang P., Baldwin J., Barry A.,
RA Bayul T., Blitshauer B., Bloom T., Blve J., Boguelavskiy L.,
RA Borowsky M., Boukhalter B., Brunache A., Butler J., Calixte N.,
RA Calvo S., Camarata J., Campo K., Chang J., Cheshatsang Y., Citroen M.,
RA Collymore A., Conside T., Cook A., Cooke P., Corum B., Cuomo C.,
RA David R., Dawoe T., Degray S., Dodge S., Dooley K., Dorje P.,
RA Dorjee K., Dorris L., Duffey N., Dupes A., Elkins T., Engels R.,
RA Erickson J., Farina A., Fero S., Ferreira P., Fischer H.,
RA Fitzgerald M., Foley K., Gage D., Galagan J., Gearin G., Gnerre S.,
RA Gnrke A., Goyette A., Graham J., Grandbois E., Gyaltsen K., Hafez N.,
RA Hagopian D., Hagos B., Hall J., Hatcher B., Heller A., Higgins H.,
RA Honan T., Horn A., Houde N., Hughes L., Hulme W., Husby E., Iliev I.,
RA Jaffe D., Jones C., Kanar M., Kanat A., Kamvasselis M., Karlsson E.,
RA Kells C., Kieu A., Kiser P., Kodira C., Kulbokas E., Labutti K.,
RA Lana D., Landers T., Leger J., Levine S., Lewis D., Lewis T.,
RA Lindblad-toh K., Liu X., Lokitsang T., Lokitsang Y., Lucien O.,
RA Lui A., Ma L.J., Mabbitt R., Macdonald J., Maclean C., Major J.,
RA Manning J., Maraballa R., Maru K., Matthews C., Mauceli E.,
RA McCarthy M., McDonough S., McGhee T., Meldrim J., Meneus L.,
RA Mesirov J., Mihalev A., Mihova T., Mikkelsen T., Mlenga V., Moru K.,
RA Mozes J., Mulrain L., Munson G., Naylor J., Nieves C., Nguyen C.,
RA Nguyen N., Nguyen T., Nicol R., Nielsen C., Nizzari M., Norbu C.,
RA Norbu N., O'donnell P., Okoawo O., O'leary S., Omotohio B.,
RA O'Neill K., Osman S., Parker S., Perrin D., Phunkhang P., Pignani B.,
RA Purcell S., Rachupka T., Ramasamy U., Rameau R., Ray V., Raymond C.,
RA Rett R., Richardson S., Rise C., Rodriguez J., Rogers J., Rogov P.,
RA Rutman M., Schupbach R., Seaman C., Settippalli S., Sharpe T.,
RA Shender J., Sherpa N., Shi J., Smirnov S., Smith C., Sougnez C.,
RA Spencer B., Stalker J., Stange-thomann N., Stavropoulos S.,
RA Stetson K., Stone C., Stone S., Stubbs M., Talamas J., Tchuinga P.,
RA Tenzing P., Tesfaye S., Theodore J., Thoulutsang Y., Topham K.,
RA Towey S., Tsamla T., Tsomo N., Vallee D., Vassiliev H.,
RA Venkataraman V., Vinson J., Vo A., Wade C., Wang S., Wangchuk T.,
RA Wangdi T., Whittaker C., Wilkinson J., Wu Y., Wyman D., Yadav S.,
RA Yang S., Yang X., Yeager S., Yee E., Young G., Zainoun J., Zembeck L.,
RA Zimmer A., Zody M., Lander E.;
RT "The genome sequence of Ustilago maydis."
RT Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL: AACP01000259; EAK83544.1; -; Genomic_DNA.
SQ SEQUENCE 488 AA; 53032 MW; 18D4BD384687A61B CRC64;

Query Match 89.6%; Score 43; DB 2; Length 488;
Best Local Similarity 87.5%; Pred. No. 5.4e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 PLPPLPXP 12
| | | | |
Db 191 PLPPLPSP 198

RESULT 27
Q52SK2_9RICK PRELIMINARY; PRT; 493 AA.
AC Q52SK2;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE VirB10.
GN Name=VirB10;
OS Wolbachia endosymbiont of Armadillidium vulgare.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Rickettsiaceae; Wolbachiae; Wolbachia.
OX NCBI_TaxID=77039;
RN [1]

```

```

RP NUCLEOTIDE SEQUENCE.
RA Felix C., Greve P., Braquart-Varnier C., Braig H.R., Martin G.;
RT "Characterization and transcriptional analysis of two gene clusters
RT for a type IV secretion machinery in feminizing intracellular symbiont
RT Wolbachia.";
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY967767; AAX86707.1; -; Genomic_DNA.
SQ SEQUENCE 493 AA; 54029 MW; 3BB89957279FA02D CRC64;

Query Match 89.6%; Score 43; DB 2; Length 493;
Best Local Similarity 87.5%; Pred. No. 5.4e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 PLPPLPXP 12
| | | | |
Db 95 PLPPLPPT 102

RESULT 28
Q4WXS5_ASFPF PRELIMINARY; PRT; 551 AA.
AC Q4WXS5;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=Afu309730;
OS Aspergillus fumigatus Af293.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=330879;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Af293;
RA Nierman W., Pain A., Anderson M.J., Wortman J., Kim H.Stanley.,
RA Arroya J., Berriman M., Abe K., Archer D.B., Bermejo C., Bennett J.,
RA Bowyer P., Chen D., Collins M., Coulson R., Davies R., Dyer P.S.,
RA Farman M., Fedorova N., Fedorova N., Feldblyum T.V., Fischer R.,
RA Fosker N., Fraser A., Garcia J.L., Garcia M.J., Goble A.,
RA Goldman G.H., Gomi K., Griffith-Jones S., Gwilliam R., Haas B.,
RA Haas H., Harris D., Horiuchi H., Huang J., Humphrey S., Jimenez J.,
RA Keller N., Khouri H., Kitamoto K., Kobayashi T., Kulkarni R.,
RA Kumagai T., Lafton A., Latge J.-P., Li W., Lord A., Lu C.,
RA Majoros W.H., May G.S., Miller B.L., Mohamoud Y., Molina M., Monod M.,
RA Mouyna I., Mulligan S., Murphy L., O'Neill S., Paulsen I.,
RA Penava M.A., Pertea M., Price C., Pritchard B.L., Quail M.A.,
RA Rabinowitsch E., Rawlins N., Rajandream M.-A., Reichard U.,
RA Renauld H., Robson G.D., Rodriguez de Cordoba S., Rodriguez-Pena J.M.,
RA Ronning C.M., Rutter S., Salzberg S.L., Sanchez M.,
RA Sanchez-Ferrero J.C., Saunders D., Seeger K., Squares R., Squares S.,
RA Takeuchi M., Tekala F., Turner G., Vazquez de Aldana C.R., Weidman J.,
RA White O., Woodward J., Yu J.-H., Fraser C., Galagan J.E., Asai K.,
RA Machida M., Hall N., Barrell B., Denning D.W.;
RT "Genomic sequence of the pathogenic and allergenic filamentous fungus
RT Aspergillus fumigatus."
RT Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL: AAHF01000002; EAL92608.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 551 AA; 61426 MW; 6743CC767983CFD3 CRC64;

Query Match 89.6%; Score 43; DB 2; Length 551;
Best Local Similarity 77.8%; Pred. No. 6.1e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 4 RPLPPLPXP 12
| | | | |
Db 242 QPLPPLPNP 250

RESULT 29

```

```

Q627E0_ORYSA
ID Q627E0_ORYSA PRELIMINARY; PRT; 554 AA.
AC Q627E0;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 03-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Putative pentatricopeptide (PPR) repeat-containing protein.
GN Name=P0020C11.29-1;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 2, PAC
clone:IP0020C11."
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP004865; BAD15828.1; -; Genomic_DNA.
DR Gramene; Q627E0; -.
DR InterPro; IPR002885; PPR.
DR InterPro; IPR011990; TPR-like_helical.
DR Pfam; PF01535; PPR; 13.
DR TIGRFAMs; TIGR00756; PPR; 12.
KW Repeat.
SQ SEQUENCE 554 AA; 60020 MW; 440B2408B8CFE044 CRC64;

Query Match 89.6%; Score 43; DB 2; Length 554;
Best Local Similarity 77.8%; Pred. No. 6.1e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 RPLPPLPXP 12
||| |||||
DB 26 RPYPLPAP 34

RESULT 30
Q55299_CRYNE
ID Q55299_CRYNE PRELIMINARY; PRT; 640 AA.
AC Q55299;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=CNB3830;
OS Cryptococcus neoformans var. neoformans B-3501A.
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Heterobasidiomycetes;
OC Tremellomycetidae; Tremellales; Tremellaceae; Filobasidiella.
OX NCBI_TaxID=283643;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=B-3501A;
RA Fung E., Hyman R.W., Rowley D., Bruno D., Miranda M., Fukushima M.,
RA Wickes B.L., Fu J., Davis R.W.;
RT "Cryptococcus neoformans serotype D sequencing."
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAEY0100038; EAL19755.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 640 AA; 67879 MW; 63F41A83A7E7E1BD CRC64;

Query Match 89.6%; Score 43; DB 2; Length 640;
Best Local Similarity 87.5%; Pred. No. 7.1e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 PLPLPLPXP 12
|||||
DB 130 PLPLPLPSP 137

```

```

RESULT 31
Q6Z0N5_ORYSA
ID Q6Z0N5_ORYSA PRELIMINARY; PRT; 745 AA.
AC Q6Z0N5;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Putative threonyl-tRNA synthetase.
GN Name=QJ1212.C09.6; Synonyms=QJ1705.A03.38;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsumoto T., Katayose Y.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsumoto T., Yamamoto K.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP005487; BAD01414.1; -; Genomic_DNA.
DR EMBL; AP003918; BAD01219.1; -; Genomic_DNA.
DR Gramene; Q6Z0N5; -.
DR GO; GO:000524; F:ATP binding; IEA.
DR GO; GO:0016874; F:ligase activity; IEA.
DR GO; GO:0004829; F:threonine-tRNA ligase activity; IEA.
DR GO; GO:0006412; P:protein biosynthesis; IEA.
DR GO; GO:0006435; P:threonyl-tRNA aminoacylation; IEA.
DR InterPro; IPR004154; HGTP anticodon.
DR InterPro; IPR004095; TGS.
DR InterPro; IPR002314; tRNA-synt 2b.
DR InterPro; IPR002320; tRNA-synt_thr.
DR InterPro; IPR006195; tRNA_ligase_II.
DR Pfam; PF03129; HGTP_anticodon; 1.
DR Pfam; PF02824; TGS; 1.
DR Pfam; PF00587; tRNA-synt 2b; 1.
DR PRINTS; PR01047; TRNASYNTHTHR.
DR TIGRFAMs; TIGR00418; thrS; 1.
DR PROSITE; PS00862; AA_TRNA_LIGASE_II; 1.
KW Aminoacyl-tRNA synthetase.
SQ SEQUENCE 745 AA; 84945 MW; EFED0C3D44974465 CRC64;

Query Match 89.6%; Score 43; DB 2; Length 745;
Best Local Similarity 77.8%; Pred. No. 8.3e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 RPLPPLPXP 12
||| |||||
DB 19 RPYPLPAP 27

RESULT 32
Q4T7J1_TETNG
ID Q4T7J1_TETNG PRELIMINARY; PRT; 851 AA.
AC Q4T7J1;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Chromosome undetermined SCAF8089, whole genome shotgun sequence.
GN ORFNames=GSTENG0005703001;
OS Tetraodon nigroviridis (green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Jaillon O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicaud S., Jaffe D., Fisher S., Fisher S., Lutfalla G., Dossat C., Segurens B.,

```

RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA Biemont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
RA Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J.,
RA Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Wolff J.N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissenbach J., Roest Crolius H.,
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RL the early vertebrate proto-karyotype.";
RN Nature 431:946-957(2004).
[2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CAABE01008089; CAP91141.1; -: Genomic DNA.
SQ SEQUENCE 851 AA; 94830 MW; A952ACD44D3388F4 CRC64;
Query Match 89.6%; Score 43; DB 2; Length 851;
Best Local Similarity 77.8%; Pred. No. 9.6e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Qy 4 RPLPPLPXP 12
Db 286 KPLPPLPPP 294
RESULT 33
Q4SDM9_TETNG
ID Q4SDM9_TETNG PRELIMINARY; PRT; 963 AA.
AC Q4SDM9;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Chromosome 10 SCAP14634, whole genome shotgun sequence.
DE (Fragment).
GN ORFNAMES=GSTENG00019959001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontoidea; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Gallion O., Aury J.M., Brunet P., Petit J.L., Stange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Doseat C., Segurens B.,
RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA Biemont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
RA Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J.,
RA Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Wolff J.N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissenbach J., Roest Crolius H.,
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RL the early vertebrate proto-karyotype.";
RN Nature 431:946-957(2004).
[2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC -!- SIMILARITY: Contains 1 C2 domain.
DR EMBL; CAABE01014634; CAG01253.1; -: Genomic_DNA.

DR InterPro: IPR000008; C2.
DR Pfam: PF00168; C2; 1.
DR PRINTS: PRO0360; C2DOMAIN.
DR SMART: SM00239; C2; 1.
DR PROSITE: PS00004; C2_DOMAIN; 1.
FT NON TER 963
SQ SEQUENCE 963 AA; 104747 MW; F496F320D2668851 CRC64;
Query Match 89.6%; Score 43; DB 2; Length 963;
Best Local Similarity 77.8%; Pred. No. 1.1e+03;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 4 RPLPPLPXP 12
Db 549 RPLPPTTP 557
RESULT 34
Q7RYUL_NEUCR
ID Q7RYUL_NEUCR PRELIMINARY; PRT; 969 AA.
AC Q7RYUL;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Predicted protein.
GN Name=NCU00367.1;
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=OR74A;
RA Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,
RA Jaffe D., FitzHugh W., Ma L.-J., Smirnov S., Purcell S., Rehan B.,
RA Elkins T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,
RA Qui D., Ianakiev P., Pedersen D., Nelson M., Washburne M.,
RA Selitrenikoff C.P., Kinsey J.A., Braun E.L., Zelter A., Schulte U.,
RA Kothe G.O., Jedd G., Mewes W., Staben C., Marcotte E., Greenberg D.,
RA Roy A., Foley K., Naylor J., Thomann N., Barrett R., Gnerre S.,
RA Kamal M., Kamvysselis M., Mauceli E., Bielke C., Rudd S., Frishman D.,
RA Kvyatova S., Rasmussen C., Metzberg R.L., Perkins D.D., Kroken S.,
RA Cystogova S., Macino G., Catchside D., Li W., Pratt R.J., Osmari S.A.,
RA Desouza C.C., Glass L., Orbach M.J., Berglund J., Voelker R.,
RA Yarden O., Plamann M., Seiler S., Dunlap J., Radford A., Aramayo R.,
RA Natvig D.O., Alex L.A., Mannhaupt G., Ebbole D.J., Freitag M.,
RA Paulsen I., Sachs M.S., Lander E.S., Nusbaum C., Birren B.,
RT "The Genome Sequence of the Filamentous Fungus Neurospora crassa.";
RL Nature 0:0-0(2003).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AABX01000723; EAA28017.1; -: Genomic DNA.
SQ SEQUENCE 969 AA; 109976 MW; 77860EC9134F1429 CRC64;
Query Match 89.6%; Score 43; DB 2; Length 969;
Best Local Similarity 87.5%; Pred. No. 1.1e+03;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 5 PLPLPXP 12
Db 574 PLPLPPTP 581
RESULT 35
Q9USX0_DROME
ID Q9USX0_DROME PRELIMINARY; PRT; 1299 AA.
AC Q9USX0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Prickle sp1e isoform.
GN Name=PK; ORFNAMES=CG12830;

OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=isoogenic dp cn bw;
 RX MEDLINE=99415814; PubMed=10485852; DOI=10.1101/gad.13.17.2315;
 RA Gubb D., Green C., Huen D., Coulson D., Johnson G., Tree D.,
 RA Collier S., Roote J.;
 RT "The balance between isoforms of the prickly LIM domain protein is
 RT critical for planar polarity in Drosophila imaginal discs";
 RL Genes Dev. 13:2315-2327(1999).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=isoogenic dp cn bw;
 RA Gubb D.C.;
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: Contains 2 LIM zinc-binding domains.
 DR EMBL; AJ243710; CAB57345.3; -; mRNA.
 DR HSSP; P04006; LIM.
 DR FlyBase; FBgn003090; CG12830.
 DR FlyBase; FBgn003090; pk.
 DR GO; GO:0005737; C:cytoplasm; TAS.
 DR GO; GO:0016020; C:membrane; NAS.
 DR GO; GO:0005515; F:protein binding; NAS.
 DR GO; GO:0042067; P:establishment of ommatidial polarity (sensu. . .; IMP.
 DR GO; GO:0045184; P:establishment of protein localization; TAS.
 DR GO; GO:0001737; P:establishment of wing hair orientation; TAS.
 DR GO; GO:0045185; P:maintenance of protein localization; TAS.
 DR InterPro; IPR001781; LIM.
 DR InterPro; IPR010442; PET.
 DR InterPro; IPR007087; Znf_C2H2.
 DR Pfam; PF00412; LIM; 2.
 DR Pfam; PF06297; PET; 1.
 DR SMART; SM00132; LIM; 3.
 DR PROSITE; PS00478; LIM DOMAIN 1; 2.
 DR PROSITE; PS00023; LIM DOMAIN 2; 3.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.
 DR LIM domain, Metal-binding; Zinc.
 SQ SEQUENCE 1299 AA; 140529 MW; 3B6D3A31717BE7DE CRC64;
 Query Match 89.6%; Score 43; DB 2; Length 1299;
 Best Local Similarity 77.8%; Pred. No. 1.5e+03;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Caps 0;
 QY 4 RPLPLPXP 12
 Db 262 RPVPPLPP 270
 ID Q3V419 DROME PRELIMINARY; PRT; 1299 AA.
 AC Q9V419
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE CG11084-PC, isoform C.
 GN Namespk; ORFNames=CG11084;
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Ananides P.G., Scher S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Mortman J.R., Vandeell M.D., Zhang Q., Chen L.X.,
 RA

RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Bortova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foslcr C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mout S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler P., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M.C., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun B.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou S., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster";
 RL Science 287:2185-2195 (2000).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=22426065; PubMed=12537568;
 RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
 RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
 RA George R.A., Hoskins R.A., Lavery T., Muzny D.M., Nelson C.R.,
 RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
 RA Svirskas R., Taber P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
 RA Weinstein G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
 RT "Finishing a whole-genome shotgun: release 3 of the Drosophila
 RT melanogaster euchromatic genome sequence";
 RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
 RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=22426070; PubMed=12537573;
 RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J.W., Svirskas R.,
 RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
 RA Ashburner M., Celniker S.E.;
 RT "The transposable elements of the Drosophila melanogaster euchromatin:
 RT a genomics perspective";
 RL Genome Biol. 3:RESEARCH0084.1-RESEARCH0084.20(2002).
 RN [4]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=22426069; PubMed=12537572;
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
 RA Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
 RA Lewis S.E.;
 RT "Annotation of the Drosophila melanogaster euchromatic genome: a
 RT systematic review";
 RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
 RN [5]
 RP NUCLEOTIDE SEQUENCE.
 RG Berkeley Drosophila Genome Project;
 RA Celniker S., Carlson J., Wan K., Pfeiffer B., Frise E., George R.,
 RA Hoskins R., Stapleton M., Pacleb J., Park S., Svirskas R., Smith E.,

RA Yu C., Rubin G.;
 RL "Drosophila melanogaster release 4 sequence."; ;
 RT Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP NUCLEOTIDE SEQUENCE.
 RG FlyBase;
 RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: Contains 2 LIM zinc-binding domains.
 DR EMBL; AE003842; AAF59281.2; -; Genomic_DNA.
 DR HSSP; P04006; LIML.
 DR Ensembl; CG11084; Drosophila melanogaster.
 DR FlyBase; FBgn0003090; CG12830.
 DR FlyBase; FBgn0003090; pk.
 DR GO; GO:0005737; C:cytoplasm; TAS.
 DR GO; GO:0016020; C:membrane; NAS.
 DR GO; GO:0005515; F:protein binding; NAS.
 DR GO; GO:0042067; P:establishment of ommitidial polarity (sensu. . .; IMP.
 DR GO; GO:0045184; P:establishment of protein localization; TAS.
 DR GO; GO:0001737; P:establishment of wing hair orientation; TAS.
 DR GO; GO:0045185; P:maintenance of protein localization; TAS.
 DR InterPro; IPR001781; LIM.
 DR InterPro; IPR010442; PTF.
 DR Pfam; PF00412; LIM; 2.
 DR Pfam; PF06297; PTF; 1.
 DR ProDom; PD000094; LIM; 3.
 DR SMART; SM00132; LIM; 3.
 DR PROSITE; PS00478; LIM DOMAIN 1; 2.
 DR PROSITE; PS0023; LIM DOMAIN 2; 3.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.
 KW LIM domain; Metal-binding; Zinc. 8BFAF1P75F352485 CRC64;
 SQ SEQUENCE 1299 AA; 140721 MW; 8BFAF1P75F352485 CRC64;

Query Match 89.6%; Score 43; DB 2; Length 1299;
 Best Local Similarity 77.8%; Pred. No. 1.5e+03;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 4 RPLPLPXP 12
 ||:|||||
 Db 262 RVPVPLPSP 270

RESULT 37
 Q4QEP3_LEIMA
 ID Q4QEP3_LEIMA PRELIMINARY; PRT; 2145 AA.
 AC Q4QEP3;
 DT 13-SEP-2005 (TrEMBLrel. 31, Created)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
 DE Hypothetical protein.
 GN ORFNames=LmjF16.1240;
 OS Leishmania major.
 OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
 OX NCBI_TaxID=5664;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=Friedlin;
 RA Peacock C.S., Murphy L., Ivens A.C., Berriman M., Blackwell J.,
 RA Smith D., Collins M., Foster N., Harris D., Oliver K., O'Neil S.,
 RA Saunders D., Seeger K., Warren T., Rajandream M., and Barrell B.G.;
 RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
 DR EMBL; CT005255; CAJ03948.1; -; Genomic_DNA.
 KW Hypothetical protein.
 SQ SEQUENCE 2145 AA; 227134 MW; EC538A1C8C088F69 CRC64;

Query Match 89.6%; Score 43; DB 2; Length 2145;
 Best Local Similarity 87.5%; Pred. No. 2.5e+03;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 PLPPLPXP 12
 |||||||
 Db 1231 PLPPLPSP 1238

RESULT 38
 Q5XWJ3_9PICO
 ID Q5XWJ3_9PICO PRELIMINARY; PRT; 2432 AA.
 AC Q5XWJ3;
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE Polyprotein.
 OS Aichi virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Kobuvirus.
 OX NCBI_TaxID=72149;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=BAY/1/03/DEU;
 RA Oh D.-Y., Hauroeder B., Schreier E.;
 RT "Isolation and characterization of Aichi virus in Germany.";
 RT Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY747174; AAV28657.1; -; Genomic_RNA.
 DR GO; GO:0019028; C:viral capsid; IEA.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0003723; F:RNA binding; IEA.
 DR GO; GO:0003724; F:RNA helicase activity; IEA.
 DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR GO; GO:0006350; P:transcription; IEA.
 DR GO; GO:0019079; P:viral genome replication; IEA.
 DR InterPro; IPR003959; AAA ATPase centr.
 DR InterPro; IPR000199; Pept_C3 piCorn.
 DR InterPro; IPR004004; Pept_Calici.
 DR InterPro; IPR001676; Rhv.
 DR InterPro; IPR000605; RNA helicase.
 DR InterPro; IPR007095; RNA_pol_DS_PS.
 DR InterPro; IPR001205; RNA_pol_P3D.
 DR InterPro; IPR007094; RNA_pol_PSVir.
 DR Pfam; PF00004; AAA; 1.
 DR Pfam; PF00548; Peptidase_C3; 1.
 DR Pfam; PF00680; RdRP_1; 1.
 DR Pfam; PF00073; Rhv; 2.
 DR Pfam; PF00910; RNA helicase; 1.
 DR PRINTS; PR00918; CALICVIRUSNS.
 KW Polyprotein.
 SQ SEQUENCE 2432 AA; 265525 MW; BE9102F34E8979E7 CRC64;
 Query Match 89.6%; Score 43; DB 2; Length 2432;
 Best Local Similarity 87.5%; Pred. No. 2.9e+03;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 PLPPLPXP 12
 |||||||
 Db 996 PLPPLPSP 1003

RESULT 39
 Q91QP4_9PICO
 ID Q91QP4_9PICO PRELIMINARY; PRT; 2432 AA.
 AC Q91QP4;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Polyprotein.
 OS Aichi virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Kobuvirus.
 OX NCBI_TaxID=72149;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC MEDLINE=21376426; PubMed=11483747;
 RX DOI=10.1128/JVI.75.17.8021-8030.2001;
 RA Sasaki J., Kusunohara Y., Maeno Y., Kobayashi N., Yamashita T.,
 RA Sakae K., Takeda N., Taniguchi K.;
 RT "Construction of an infectious cDNA clone of Aichi virus (a new member

```
RT of the family Picornaviridae) and mutational analysis of a stem-loop
RT structure at the 5' end of the genome."
RL J. Virol. 75:8021-8030(2001).
DR EMBL; AB040749; BAB62889.1; -; Genomic_RNA.
DR HSSP; P03300; IRDR.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003724; F:RNA helicase activity; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR InterPro; IPR003959; AAA_ATPase_Centr.
DR InterPro; IPR000199; Pept_C3_picorn.
DR InterPro; IPR004004; Pept_Calici.
DR InterPro; IPR001676; Rhv.
DR InterPro; IPR000605; RNA_helicase.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR001205; RNA_pol_P3D.
DR InterPro; IPR007094; RNA_pol_P5vir.
DR Pfam; PF00004; AAA; 1.
DR Pfam; PF00548; Peptidase_C3; 1.
DR Pfam; PF00680; RdRP_1; 1.
DR Pfam; PF00073; Rhv_2.
DR PRINTS; PR00918; CALICVIRUSNS.
KW Polyprotein.
SQ SEQUENCE 2432 AA; 265456 MW; E451BABAD3CD1460 CRC64;

Query Match 89.6%; Score 43; DB 2; Length 2432;
Best Local Similarity 87.5%; Pred. No. 2.9e+03;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 PLPPLXP 12
Db 996 PLPPLPTP 1003

RESULT 40
Q91464_9PICO PRELIMINARY; PRT; 2433 AA.
AC Q91464;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Aichi virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Kobuvirus.
OX NCBI_TaxID=72149;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=846/88;
RX MEDLINE=98406258; PubMed=9733894;
RA Yamashita T., Sakae K., Tazuki H., Suzuki Y., Ishikawa N., Takeda N.,
RA Miyamura T., Yamazaki S.
RT "Complete nucleotide sequence and genetic organization of Aichi virus,
RT a distinct member of the Picornaviridae associated with acute
RT gastroenteritis in humans."
RL J. Virol. 72:8408-8412(1998).
DR EMBL; AB010145; BAA31356.1; -; Genomic_RNA.
DR HSSP; P03300; IRDR.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003724; F:RNA helicase activity; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR InterPro; IPR003959; AAA_ATPase_Centr.
DR InterPro; IPR000199; Pept_C3_picorn.
```

```
DR InterPro; IPR004004; Pept_Calici.
DR InterPro; IPR001676; Rhv.
DR InterPro; IPR000605; RNA_helicase.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR001205; RNA_pol_P3D.
DR InterPro; IPR007094; RNA_pol_P5vir.
DR Pfam; PF00004; AAA; 1.
DR Pfam; PF00548; Peptidase_C3; 1.
DR Pfam; PF00680; RdRP_1; 1.
DR Pfam; PF00073; Rhv_2.
DR PRINTS; PR00918; CALICVIRUSNS.
KW Polyprotein.
FT CHAIN 171 540 VP0.
FT CHAIN 541 764 VP3.
FT CHAIN 765 1042 VP1.
FT CHAIN 1043 1153 2A.
FT CHAIN 1154 1318 2B.
FT CHAIN 1319 1653 2C.
FT CHAIN 1654 1748 3A.
FT CHAIN 1749 1775 3B.
FT CHAIN 1776 1965 3C.
FT CHAIN 1966 2433 3D.
SQ SEQUENCE 2433 AA; 265526 MW; 44EFBBE1B76D6188 CRC64;

Query Match 89.6%; Score 43; DB 2; Length 2433;
Best Local Similarity 87.5%; Pred. No. 2.9e+03;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 PLPPLXP 12
Db 997 PLPPLPTP 1004

RESULT 41
Q4TWI0_9PICO PRELIMINARY; PRT; 2433 AA.
AC Q4TWI0;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Polyprotein.
OS Aichi virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Kobuvirus.
OX NCBI_TaxID=72149;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Goiania/GO/03/01/Brazil;
RA Silva P.A., Cardoso D.D.P., Schreier E.;
RT "Molecular characterization of aichi virus from Brazil."
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; DQ028632; AAY46271.1; -; Genomic_RNA.
KW Polyprotein.
SQ SEQUENCE 2433 AA; 266277 MW; 298F73C2A60AE53C CRC64;

Query Match 89.6%; Score 43; DB 2; Length 2433;
Best Local Similarity 87.5%; Pred. No. 2.9e+03;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 PLPPLXP 12
Db 997 PLPPLPTP 1004

RESULT 42
Q5WLC4_BACSK PRELIMINARY; PRT; 80 AA.
ID Q5WLC4_BACSK
AC Q5WLC4;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Inner spore coat protein D.
```

```

GN OrderedLocusNames=ABC0289;
OS Bacillus clausii (strain KSM-K16).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=66692;
RN [1]
RC NUCLEOTIDE SEQUENCE.
RA Takaki Y., Kageyama Y., Shimamura S., Suzuki H., Nishi S., Hatada Y.,
RA Kawai S., Ito S., Horikoshi K.; of the alkaliphilic Bacillus clausii
RT "The complete genome sequence of the alkaliphilic Bacillus clausii
KSM-K16.";
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP006627; BAD62831.1; -; Genomic_DNA.
KW Capsid protein; Complete proteome.
SQ SEQUENCE 80 AA; 9185 MW; 908297ED8E2862E3 CRC64;

Query Match 87.5%; Score 42; DB 2; Length 80;
Best Local Similarity 77.8%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 RLPLPLXP 12
Db 15 RLPLPAPRP 23

RESULT 43
Q4H8A2_9DEIO PRELIMINARY; PRT; 106 AA.
AC Q4H8A2;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=DgeodRAFT_0541;
OS Deinococcus geothermalis DSM 11300.
OC Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales;
OC Deinococcaceae; Deinococcus.
OX NCBI_TaxID=319795;
RN [1]
RC NUCLEOTIDE SEQUENCE.
RA US DOE Joint Genome Institute (JGI-PGF);
RA Copeland A., Lucas S., Lapidus A., Barry K., Detter C., Glavina T.,
RA Hammon N., Israni S., Pittluck S., Richardson P.;
RT "Sequencing of the draft genome assembly of Deinococcus geothermalis
DSM 11300.";
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
RN [2]
RC NUCLEOTIDE SEQUENCE..
RA US DOE Joint Genome Institute (JGI-ORNL);
RA Larimer F., Land M.;
RT "Annotation of the draft genome assembly of Deinococcus geothermalis
DSM 11300.";
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAHE0100007; EAL82441.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 106 AA; 11553 MW; 0C49351FA56D3310 CRC64;

Query Match 87.5%; Score 42; DB 2; Length 106;
Best Local Similarity 87.5%; Pred. No. 1.5e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 PLPLPLXP 12
Db 8 PLPLPLPL 15

RESULT 44
Q5AGI6_CANAL PRELIMINARY; PRT; 134 AA.
AC Q5AGI6;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Hypothetical protein.
GN ORFNames=CaO19.11804;
OS Candida albicans SC5314.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=237561;
RN [1]
RC NUCLEOTIDE SEQUENCE.
RA Jones T., Federspiel N.A., Chibana H., Dungan J., Kalman S., Magee P.T.,
RA Magee B.B., Newport G., Thorstenson Y.R., Agabian N., Magee P.T.,
RA Davis R.W., Scherer S.;
RT "The diploid genome sequence of Candida albicans.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:7329-7334(2004).
RN [2]
RC NUCLEOTIDE SEQUENCE.
RA Dungan J., Kuo A., Newport G., Lan C.-Y., Iijima C., Adegbola O.,
RA Roberts J., Persson K., Donnelly S., Favoreto S., Tzung K.-W.,
RA Jones T., Scherer S., Agabian N.;
RT "Annotation of the genome of Candida albicans.";
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AACQ01000021; EAL01934.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 134 AA; 14368 MW; 06CB8DD0EF305BB CRC64;

Query Match 87.5%; Score 42; DB 2; Length 134;
Best Local Similarity 87.5%; Pred. No. 1.9e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 PLPLPLXP 12
Db 89 PLPLPLP 96

RESULT 45
Q5Z6V2_ORYSA PRELIMINARY; PRT; 144 AA.
AC Q5Z6V2;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein P0551A03.21.
GN Names=P0551A03.21;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RC NUCLEOTIDE SEQUENCE..
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 6, PAC
clone:P0551A03.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP004818; BAD54317.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 144 AA; 15117 MW; 008F8B457FB71A6B CRC64;

Query Match 87.5%; Score 42; DB 2; Length 144;
Best Local Similarity 87.5%; Pred. No. 2e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 PLPLPLXP 12

```

```

Db          ||||| |
71 PLPPLPEP 78

RESULT 46
Q52B26_ORYSA PRELIMINARY; PRT; 150 AA.
AC Q52B26;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Hypothetical protein OJ111_G12.27 (Hypothetical protein
DE B1043F02.2).
GN Names:OJ111_G12.27; Synonyms:B1045F02.2;
OS Oryza sativa (Japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzoideae; Oryza.
OX NCBI_TaxID=39947;
RN NUCLEOTIDE SEQUENCE.
RP Sasaki T., Matsumoto T., Yamamoto K., Sakata K., Baba T., Katayose Y.,
RA Wu J., Niimura Y., Cheng Z., Nagamura Y., Antonio B.A., Kanamori H.,
RA Okochiwa S., Masukawa M., Arikawa K., Chiden Y., Hayashi M.,
RA Hosokawa M., Ando T., Aoki H., Arita K., Hamada M., Harada C.,
RA Hijishita S., Honda M., Ichikawa Y., Idonuma A., Iijima M., Ikeda M.,
RA Ikono M., Itoh S., Itoh Y., Itoh Y., Iwabuchi A., Kamiya K.,
RA Karasawa W., Katagiri S., Kikuta A., Kobayashi N., Kono I.,
RA Machita K., Maehara T., Mizuno H., Mizubayashi T., Mukai Y.,
RA Nagasaki H., Nakashima M., Nakama Y., Nakamichi Y., Nakamura M.,
RA Namiki N., Negishi M., Ohta I., Ono N., Saji S., Sakai K., Shibata M.,
RA Shimokawa T., Shomura A., Song J., Takazaki Y., Terasawa K., Tsuji K.,
RA Waki K., Yanagata H., Yanane H., Yoshiki S., Yoshihara R., Yukawa K.,
RA Zhong H., Iwama H., Endo T., Ito H., Hahn J.H., Kim H.I., Eun M.Y.,
RA Yano M., Jiang J., Gojobori T.;
RT "The genome sequence and structure of rice chromosome 1.";
RL Nature 420:312-316(2002).
DR EMBL; AP003327; BAD53216.1; -; Genomic DNA.
DR EMBL; AP003329; BAD53080.1; -; Genomic DNA.
KW Hypothetical protein.
SQ
SEQUENCE 150 AA; 16137 MW; C85985612A904DFF CRC64;

Query Match 87.5%; Score 42; DB 2; Length 150;
Best Local Similarity 87.5%; Pred. No. 2.1e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 PLPPLXP 12
Db 18 PLPPLPRP 25

RESULT 47
Q8BG6_MOUSE PRELIMINARY; PRT; 151 AA.
AC Q8BG6;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Mus musculus 10 days neonate cerebellum cDNA, RIKEN full-length
DE enriched library, clone:B930008A12 product:similar to PP2CH (Mus
DE musculus 0 day neonate eyeball cDNA, RIKEN full-length enriched
DE library, clone:E130208E03 product:similar to PP2CH).
GN Names:Ppm1e;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN NUCLEOTIDE SEQUENCE.
RP Carninci P., Hayashizaki Y.;
RC STRAIN=C57BL/6J; TISSUE=Cerebellum, and Eyeball;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;

"High-efficiency full-length cDNA cloning.";
Meth. Enzymol. 303:19-44(1999).
[2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum, and Eyeball;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadosa K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Ronald M.F.,
RA Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombauts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seva T., Shibata Y., Storch K.-F.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kontecki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
[3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum, and Eyeball;
RX MEDLINE=22354683; PubMed=1246851; DOI=10.1038/nature01266;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nozaki A., Schonbach C., Gojobori T.,
RA Baldarelli R., Hill D.P., Bull C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasaki Y., Kedzierski R.M., King B.L.,
RA Kongaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wyshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavalan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
[4]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum, and Eyeball;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
[5]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum, and Eyeball;
RX MEDLINE=2030913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,

```

RA Konno H., Akiyama J., Nishi K., Kitsuunai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RA "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771 (2000).
RN [6]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum, and Eyeball;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Kori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Kato H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsuura T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toyota T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AK046962; BAC32927.1; -; mRNA.
DR EMBL: AK053696; BAC35479.1; -; mRNA.
DR MGI: 2444096; Ppmie.
SQ SEQUENCE 151 AA; 16372 MW; 868788C91FDD452E CRC64;
Query Match 87.5%; Score 42; DB 2; Length 151;
Best Local Similarity 87.5%; Pred. No. 2.1e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 5 PLPPLXP 12
DB 120 PLPPLXP 127
RESULT 48
ID Q8PNK3_XANAC PRELIMINARY; PRT; 176 AA.
AC Q8PNK3;
DT 01-OCT-2002 (T-EMBLrel. 22, Created)
DT 01-OCT-2002 (T-EMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
DE Hypothetical protein XAC1056.
GN OrderedLocusNames=XAC1056;
OS Xanthomonas axonopodis (pv. citri).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=92829;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=306 / ATCC 13902 / XV 101;
RX MEDLINE=22022145; PubMed=12024217; DOI=10.1038/417459a;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Almeida R.B., Monteiro-Vitorello C.B., Van Sluys M.A.,
RA Almeida N.F., Jr., Alves L.M.C., do Amaral A.M., Bertolini M.C.,
RA Camargo L.E.A., Camarotte G., Cannavan F., Cardozo J., Chambergo F.,
RA Ciapina L.P., Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R.,
RA El-Dorry H., Faria J.B., Ferreira A.J.S., Ferreira R.C.C.,
RA Ferro M.I.T., Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White P.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities."
RL Nature 417:459-463 (2002).

DR EMBL: AE011735; AAM35936.1; -; Genomic DNA.
KW Complete proteome; Hypothetical protein
SQ SEQUENCE 176 AA; 18964 MW; 14242128E08E56FF CRC64;
Query Match 87.5%; Score 42; DB 2; Length 176;
Best Local Similarity 87.5%; Pred. No. 2.5e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 5 PLPPLXP 12
DB 50 PLPPLXP 57
RESULT 49
ID Q9D7T0_MOUSE PRELIMINARY; PRT; 188 AA.
AC Q9D7T0;
DT 01-JUN-2001 (T-EMBLrel. 17, Created)
DT 01-JUN-2001 (T-EMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (T-EMBLrel. 17, Last annotation update)
DE Mus musculus adult male stomach cDNA, RIKEN full-length enriched
DE library, clone:2210411K1 product:hypothetical protein, full insert
DE sequence.
GN Names=2210411K1; Rik;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muroidae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Stomach;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RA "High-efficiency full-length cDNA cloning."
RL Meth. Enzymol. 303:19-44 (1999).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Stomach;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadoya K., Matsuura H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schram L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaiji H., Kontecki S.,
RA Hayashizaki Y.;
RL "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690 (2001).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Stomach;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573 (2002).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Stomach;
RX MEDLINE=2049374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,

RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes."
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Stomach;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagao S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamanoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaki S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Onodera Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kita A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multicapillary sequencer."
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Stomach;
RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,
RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
RA Inotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,
RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,
RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
RA Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK008900; BAB25960.1; -; mRNA.
DR Ensembl; ENSMUSG0000030431; Mus musculus.
DR MGI; MGI:1922935; 221041klrik.
KW Hypothetical protein.
SQ SEQUENCE 188 AA; 19274 MW; 4B6F8DAD2EDCB327 CRC64;

Query Match 87.5%; Score 42; DB 2; Length 188;
Best Local Similarity 77.8%; Pred. No. 2.7e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 RPLPLPXP 12
|||
DB 147 RPSPLPAP 155

RESULT 50
ID 001621 CAEEL PRELIMINARY; PRT; 198 AA.
AC 001621
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein W08A12.3.
GN ORFNames=W08A12.3;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RG The C. elegans sequencing consortium;
RT "Genome sequence of the nematode C. elegans: a platform for
investigating biology."
RL Science 282:2012-2018(1998).
DR EMBL; U97010; BAB52319.1; -; Genomic_DNA.
DR FIR; E89008; E89008.
DR Ensembl; W08A12.3; Caenorhabditis elegans.
DR WormBase; WBGene00021080; W08A12.3.

DR WormPep; W08A12.3; CE14700.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 198 AA; 21928 MW; 7CFD2E55D123FCBF CRC64;
Query Match 87.5%; Score 42; DB 2; Length 198;
Best Local Similarity 87.5%; Pred. No. 2.8e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 5 PLPPLPXP 12
|||
DB 98 PLPPLPXP 105
Search completed: April 6, 2006, 09:38:32
Job time : 180.895 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 6, 2006, 09:22:55 ; Search time 134.211 Seconds

(without alignments)
39.286 Million cell updates/sec

Title: US-10-632-388-312

Perfect score: 67

Sequence: 1 PVPVPPPIPX 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Database :

1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*
9: Geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	67	100.0	10	2 AAR93314	AAR93314 ABL prote
2	67	100.0	10	6 ADA08234	ADA08234 Human ABL
3	67	100.0	12	2 AAR93369	AAR93369 ABL prote
4	67	100.0	12	2 AAR93371	AAR93371 ABL prote
5	67	100.0	12	3 AAB17256	AAB17256 SH3 antag
6	67	100.0	12	5 ABB73249	ABB73249 Src homol
7	67	100.0	12	7 ADJ73403	ADJ73403 SH3 antag
8	67	100.0	12	8 ADJ53037	ADJ53037 CH1 delet
9	67	100.0	12	8 ADJ51998	ADJ51998 CH1 delet
10	66	98.5	12	2 AAR93373	AAR93373 ABL prote
11	66	98.5	12	2 AAR93372	AAR93372 ABL prote
12	66	98.5	12	2 AAR93370	AAR93370 ABL prote
13	66	98.5	12	3 AAB17257	AAB17257 SH3 antag
14	66	98.5	12	5 ABB73250	ABB73250 Src homol
15	66	98.5	12	7 ADJ73404	ADJ73404 SH3 antag
16	66	98.5	12	8 ADJ53038	ADJ53038 CH1 delet
17	66	98.5	12	8 ADJ51999	ADJ51999 CH1 delet
18	60	89.6	10	2 AAR93315	AAR93315 ABL prote
19	59	88.1	10	2 AAR93318	AAR93318 ABL prote
20	59	88.1	10	2 AAR93319	AAR93319 ABL prote
21	59	88.1	11	2 AAR93374	AAR93374 ABL prote
22	59	88.1	12	2 AAR93376	AAR93376 ABL prote
23	58	86.6	10	2 AAR93317	AAR93317 ABL prote
24	58	86.6	10	2 AAR93323	AAR93323 ABL prote

25	58	86.6	10	4 AAB98984	AAB98984 Yeast ORF
26	58	86.6	10	7 ADK70159	ADK70159 Human SH3
27	58	86.6	11	6 ABG75623	ABG75623 Human ABL
28	58	86.6	40	3 AAB12074	AAB12074 SH3 bindi
29	58	86.6	40	7 ADE84647	ADE84647 3BP2 SH3
30	58	86.6	467	8 ADQ67211	ADQ67211 Novel hum
31	58	86.6	492	8 ADQ6504	ADQ6504 T cell ac
32	58	86.6	536	8 ADQ66376	ADQ66376 Novel hum
33	58	86.6	561	7 ADC71301	ADC71301 Human col
34	58	86.6	561	7 ADK70062	ADK70062 Mutant hu
35	58	86.6	561	7 ADK70058	ADK70058 Wild type
36	58	86.6	561	8 ADQ96506	ADQ96506 T cell ac
37	57	85.1	50	4 ABB19592	ABB19592 Novel hum
38	57	85.1	94	4 ABB71558	ABB71558 Drosophil
39	57	85.1	98	4 AAO02495	AAO02495 Human pol
40	57	85.1	118	4 ABG16489	ABG16489 Novel hum
41	57	85.1	142	7 ABO76127	ABO76127 Pseudomon
42	57	85.1	182	4 AAB80893	AAB80893 Human ATF
43	57	85.1	216	4 AAB80892	AAB80892 Human ATF
44	57	85.1	216	4 ABB50151	ABB50151 Human tra
45	57	85.1	227	4 ABB19594	ABB19594 Novel hum
46	57	85.1	282	4 AAB80896	AAB80896 Human ATF
47	57	85.1	282	4 AAB95859	AAB95859 Human pro
48	57	85.1	282	5 ABB81169	ABB81169 Human act
49	57	85.1	282	7 ADE60923	ADE60923 Human Pro
50	57	85.1	282	8 ADP12591	ADP12591 Protein e
51	57	85.1	282	8 ADQ20824	ADQ20824 Human sof
52	57	85.1	282	8 ADM14371	ADM14371 Human NF-
53	57	85.1	282	8 ABM82139	ABM82139 Tumour-as
54	57	85.1	308	6 ABR41338	ABR41338 Human DIT
55	57	85.1	315	8 ADT57494	ADT57494 Plant pol
56	57	85.1	397	8 ADQ97684	ADQ97684 Mouse can
57	57	85.1	417	4 ABG12837	ABG12837 Novel hum
58	57	85.1	557	8 ADY24937	ADY24937 Plant ful
59	57	85.1	783	2 AAW37151	AAW37151 Mouse neu
60	57	85.1	787	2 AAW37152	AAW37152 Mouse neu
61	57	85.1	801	8 ADQ97686	ADQ97686 Mouse can
62	57	85.1	802	2 AAW37153	AAW37153 Mouse neu
63	57	85.1	802	4 AAU09139	AAU09139 Mammalian
64	57	85.1	1134	5 AAE24341	AAE24341 Human lun
65	57	85.1	1134	7 ADF74123	ADF74123 Human nov
66	56	83.6	64	4 AAO00314	AAO00314 Human pol
67	56	83.6	104	4 AAM15554	AAM15554 Peptide #
68	56	83.6	104	4 ABB34560	ABB34560 Peptide #
69	56	83.6	104	4 AAM28043	AAM28043 Peptide #
70	56	83.6	104	4 ABB29387	ABB29387 Peptide #
71	56	83.6	104	4 ABB19968	ABB19968 Protein #
72	56	83.6	104	4 AAM67743	AAM67743 Human bon
73	56	83.6	104	4 AAM55347	AAM55347 Human bra
74	56	83.6	104	4 ABG49381	ABG49381 Human liv
75	56	83.6	104	4 AAM03303	AAM03303 Peptide #
76	56	83.6	104	5 ABG37302	ABG37302 Human pep
77	56	83.6	113	5 ADK34538	ADK34538 Novel hum
78	56	83.6	156	4 ABG14149	ABG14149 Novel hum
79	56	83.6	280	8 ADY11773	ADY11773 Plant ful
80	56	83.6	293	8 ADM97712	ADM97712 Human MNA
81	56	83.6	362	3 AAB42278	AAB42278 Human ORF
82	56	83.6	368	8 ADM97711	ADM97711 Human MNA
83	56	83.6	386	6 ABR47582	ABR47582 Breast ca
84	56	83.6	393	2 AAW37149	AAW37149 Murine En
85	56	83.6	393	4 AAU09148	AAU09148 Ena/VASP-
86	56	83.6	393	7 ADE62912	ADE62912 Rat Prote
87	56	83.6	393	7 ADD48812	ADD48812 Rat Prote
88	56	83.6	393	8 ADG46666	ADG46666 Mouse ena
89	56	83.6	393	8 ADJ14191	ADJ14191 Mouse ena
90	56	83.6	418	6 ABR47581	ABR47581 Breast ca
91	56	83.6	418	7 ADE62914	ADE62914 Human Pro
92	56	83.6	418	7 ADD48814	ADD48814 Human Pro
93	56	83.6	418	8 ADP55192	ADP55192 Human PRO
94	56	83.6	418	8 ADP24828	ADP24828 PRO polyp
95	56	83.6	418	9 ADY17279	ADY17279 PRO polyp
96	56	83.6	418	9 ADY70308	ADY70308 Human bet
97	56	83.6	456	4 AAU17420	AAU17420 Novel sig

98	56	83.6	456	7	ADB94128	ADB94128 Human nov	171	55	82.1	267	8	ADJ63977	Adj63977 Human PAS
99	56	83.6	487	4	ABG21207	ABG21207 Novel hum	172	55	82.1	267	8	ADL71817	Adl71817 Human PAS
100	56	83.6	536	8	ADM97710	Adm97710 Human MNA	173	55	82.1	267	8	ADN07588	Adn07588 Human PAS
101	56	83.6	582	4	ABB59603	Abb59603 Drosophil	174	55	82.1	268	2	AAW48953	Aaw48953 Non-cleav
102	56	83.6	745	5	RAE19569	Rae19569 Human cyt	175	55	82.1	271	2	AAW28596	Aaw28596 Fas ligan
103	56	83.6	777	5	ABP41839	Abp41839 Human ova	176	55	82.1	277	2	AAW04372	Aaw04372 Human PAS
104	56	83.6	880	9	ADX83085	Adx83085 Human PEL	177	55	82.1	277	2	AAW28595	Aaw28595 Fas ligan
105	56	83.6	905	2	AAW31186	Aaw31186 Human p16	178	55	82.1	278	6	ABR82203	Abr82203 Chimeric
106	56	83.6	945	4	ABB60000	Abb60000 Drosophil	179	55	82.1	280	8	ADJ36211	Adj36211 Self-coal
107	56	83.6	983	8	ADM87342	Adm87342 Human pro	180	55	82.1	281	2	AAR77281	Aar77281 Human PAS
108	56	83.6	1021	7	ADJ68507	Adj68507 Human hea	181	55	82.1	281	2	AAR79097	Aar79097 Human PAS
109	56	83.6	1130	7	ADC99065	Adc99065 Human KPP	182	55	82.1	281	2	AAR88356	Aar88356 Human PAS
110	56	83.6	1130	8	ADM97695	Adm97695 Human MNA	183	55	82.1	281	2	AAR98104	Aar98104 Human PAS
111	56	83.6	1130	8	ABM81812	Abm81812 Tumour-as	184	55	82.1	281	2	AAW27143	Aaw27143 Human PAS
112	56	83.6	1130	9	ADK83072	Adk83072 Proline,	185	55	82.1	281	2	AAW75959	Aaw75959 Human PAS
113	56	83.6	1135	2	AAW31185	Aaw31185 Human p16	186	55	82.1	281	2	AAW49105	Aaw49105 Fas ligan
114	56	83.6	1251	7	ADB09332	Adb09332 Novel pro	187	55	82.1	281	2	AAW98071	Aaw98071 Human PAS
115	56	83.6	1251	8	ADM87788	Adm87788 Human EST	188	55	82.1	281	2	AAW04373	Aaw04373 Human PAS
116	56	83.6	1493	4	ABB69228	Abb69228 Drosophil	189	55	82.1	281	2	AAW95041	Aaw95041 Human PAS
117	56	83.6	1953	6	ABR53597	Abr53597 Protein s	190	55	82.1	281	2	AAW28594	Aaw28594 Wild type
118	56	83.6	1953	7	ADK63514	Adk63514 Disease t	191	55	82.1	281	3	AAW87580	Aaw87580 Human PAS
119	56	83.6	1953	8	ADN19292	Adn19292 Bacterial	192	55	82.1	281	3	AAW87577	Aaw87577 Human PAS
120	56	83.6	1958	2	AAR60620	Aar60620 Protein f	193	55	82.1	281	3	AAW87569	Aaw87569 Human PAS
121	55	82.1	10	4	AAB86131	Aab86131 Proline-r	194	55	82.1	281	3	AAW87575	Aaw87575 Human PAS
122	55	82.1	10	4	AAB86149	Aab86149 Proline-r	195	55	82.1	281	3	AAW87576	Aaw87576 Human PAS
123	55	82.1	10	4	AAB70933	Aab70933 Polyoma v	196	55	82.1	281	3	AAW87579	Aaw87579 Human PAS
124	55	82.1	13	4	AAB70928	Aab70928 Polyoma v	197	55	82.1	281	3	AAW87574	Aaw87574 Human PAS
125	55	82.1	14	2	AAW38060	Aaw38060 PPGPY mot	198	55	82.1	281	3	AAW87582	Aaw87582 Human PAS
126	55	82.1	14	2	ADB49250	Adb49250 Biotinyla	199	55	82.1	281	3	AAW87578	Aaw87578 Human PAS
127	55	82.1	18	2	AAW38923	Aaw38923 Peptide r	200	55	82.1	281	3	AAW87581	Aaw87581 Human PAS
128	55	82.1	18	2	AAW39010	Aaw39010 Peptide r	201	55	82.1	281	3	AAU19342	Aau19342 Amino aci
129	55	82.1	20	2	AAW38961	Aaw38961 Peptide r	202	55	82.1	281	4	AAU04556	Aau04556 Human PAS
130	55	82.1	28	2	AAW41643	Aaw41643 Mammalian	203	55	82.1	281	5	ABB80000	Abb80000 Human PAS
131	55	82.1	29	6	AAE30902	Aae30902 53BP2 pro	204	55	82.1	281	5	ABG96462	Abg96462 Human PAS
132	55	82.1	46	4	ABG12808	Abg12808 Novel hum	205	55	82.1	281	5	ABG05000	Abg05000 Human PAS
133	55	82.1	59	3	AAO20396	Aao20396 Human sec	206	55	82.1	281	5	ABB81652	Abb81652 Human PAS
134	55	82.1	85	4	AAO20336	Aao20336 Human pol	207	55	82.1	281	5	ABE29153	Abe29153 Wild-type
135	55	82.1	85	4	AAO04573	Aao04573 Human pol	208	55	82.1	281	6	ABR42309	Abr42309 Human PAS
136	55	82.1	93	8	ADR94152	Adr94152 Novel S.	209	55	82.1	281	6	ABP60549	Abp60549 Human tum
137	55	82.1	93	9	AEA58022	Aea58022 Streptoco	210	55	82.1	281	7	ADC35194	Adc35194 Human TNF
138	55	82.1	102	5	ABP01651	Abp01651 Human ORF	211	55	82.1	281	7	ADF44995	Adf44995 Human PAS
139	55	82.1	121	8	ADR10352	Adr10352 Human pro	212	55	82.1	281	7	ABW02272	Abw02272 Human PAS
140	55	82.1	127	5	AAE29154	Aae29154 Human non	213	55	82.1	281	7	ADG14422	Adg14422 Human PAS
141	55	82.1	137	3	AAE18806	Aae18806 Arabidops	214	55	82.1	281	7	ADJ74024	Adj74024 Human PAS
142	55	82.1	138	3	AAE18805	Aae18805 Arabidops	215	55	82.1	281	7	ADJ92614	Adj92614 Human PAS
143	55	82.1	139	4	ABBA1888	Abba1888 Peptide #	216	55	82.1	281	7	ADL17694	Adl17694 Human PAS
144	55	82.1	139	4	AAW35687	Aaw35687 Peptide #	217	55	82.1	281	8	ADF90414	Adf90414 Human PAS
145	55	82.1	139	4	AAW75582	Aaw75582 Human bon	218	55	82.1	281	8	ADJ56767	Adj56767 Human PAS
146	55	82.1	139	4	AAW62764	Aaw62764 Human bra	219	55	82.1	281	8	ADL27664	Adl27664 Human PAS
147	55	82.1	139	4	ABG57324	Abg57324 Human liv	220	55	82.1	281	8	ADL23577	Adl23577 Human PAS
148	55	82.1	139	5	ABG45112	Abg45112 Human pep	221	55	82.1	281	8	ADK19674	Adk19674 Chicken a
149	55	82.1	158	4	AAW5985	Aaw5985 Human sec	222	55	82.1	281	8	ADK19673	Adk19673 Human PAS
150	55	82.1	158	4	AAW5986	Aaw5986 Human sec	223	55	82.1	281	8	ADM53436	Adm53436 Human PAS
151	55	82.1	166	8	ADY05758	Ady05758 Plant ful	224	55	82.1	281	8	ADP12464	Adp12464 Protein e
152	55	82.1	166	8	ADX90309	Adx90309 Plant ful	225	55	82.1	281	8	ADO19786	Ado19786 Human PRO
153	55	82.1	172	7	ABM87038	Abm87038 Rice abio	226	55	82.1	281	8	ADO19800	Ado19800 Human PRO
154	55	82.1	173	7	ABO75054	Abw75054 Pseudomon	227	55	82.1	281	8	ADP26984	Adp26984 Human PAS
155	55	82.1	192	4	ABA48292	Abw48292 Human ZP4	228	55	82.1	281	8	ADS87994	Ads87994 Tumour tr
156	55	82.1	192	8	ADO62766	Ado62766 Transcrip	229	55	82.1	281	8	ADP56004	Adp56004 Human PRO
157	55	82.1	197	3	AAE83079	Aae83079 F-box pro	230	55	82.1	281	8	ADP56004	Adp56004 Human PRO
158	55	82.1	197	5	AAO22465	Aao22465 Human F-b	231	55	82.1	281	8	ADR44815	Adr44815 Fas ligan
159	55	82.1	197	9	ADY62389	Ady62389 Human F-b	232	55	82.1	281	8	ADR47163	Adr47163 Human PAS
160	55	82.1	234	3	AAW32786	Aaw32786 Eucalyptu	233	55	82.1	281	8	ADT78394	Adt78394 Human PAS
161	55	82.1	236	4	AAW39701	Aaw39701 Human pol	234	55	82.1	281	9	ADY19512	Ady19512 PRO poly
162	55	82.1	239	4	ABB70310	Abb70310 Drosophil	235	55	82.1	281	9	ADY16498	Ady16498 PRO poly
163	55	82.1	248	9	ADW17139	Adw17139 Eucalyptu	236	55	82.1	281	9	ADZ14447	Adz14447 Human PAS
164	55	82.1	255	4	AAW86143	Aaw86143 A. victor	237	55	82.1	288	8	ADK00625	Adk00625 HOMO prot
165	55	82.1	255	4	AAW70926	Aaw70926 GFP-PLP c	238	55	82.1	303	8	ADI43195	Adi43195 HOMO tra
166	55	82.1	258	2	AAW04371	Aaw04371 Human PAS	239	55	82.1	303	8	ADO03266	Ado03266 Thalecres
167	55	82.1	260	7	ADG14424	Adg14424 Human PAS	240	55	82.1	310	6	ADA55489	Ada55489 Human pro
168	55	82.1	260	7	ADJ69950	Adj69950 Human hea	241	55	82.1	311	2	AAW17387	Aaw17387 Human DNA
169	55	82.1	261	2	AAW28597	Aaw28597 Fas ligan	242	55	82.1	323	3	ADM05940	Adm05940 Human pro
170	55	82.1	265	2	AAW48954	Aaw48954 Non-cleav	243	55	82.1	326	3	AAG43482	Aag43482 Arabidops

244	55	82.1	326	3	AAG24964	Aag24964 Arabidops	317	55	82.1	699	5	ABG32447	Abg32447 Human SR
245	55	82.1	328	7	ADD26784	Add26784 Human adi	318	55	82.1	699	7	ADB64660	Adb64660 Human pro
246	55	82.1	329	8	ADJ36255	Adj36255 Self-coal	319	55	82.1	699	7	ADJ37689	Adj37689 Human kin
247	55	82.1	360	4	AAB93436	Aab93436 Human pro	320	55	82.1	699	8	ADR31218	Adr31218 Human SRP
248	55	82.1	360	8	ADN05477	Adn05477 Antipsori	321	55	82.1	706	7	ADM05234	Adm05234 Human pro
249	55	82.1	360	8	ADN05382	Adn05382 Antipsori	322	55	82.1	715	8	ADX91718	Adx91718 Plant ful
250	55	82.1	360	9	ADY18819	Ady18819 PRO poly	323	55	82.1	737	9	ADV99974	Adv99974 Human PEM
251	55	82.1	360	9	ADY79929	Ady79929 Amin aci	324	55	82.1	717	9	ADV99972	Adv99972 Human PEM
252	55	82.1	362	4	ABG16456	Abg16456 Novel hum	325	55	82.1	749	7	ADB64468	Adb64468 Human pro
253	55	82.1	375	7	ADE09101	Ade09101 Novel pro	326	55	82.1	753	7	ADG31678	Adg31678 Human pro
254	55	82.1	375	9	AAD40616	Aad40616 Novel hum	327	55	82.1	798	5	AAU93051	Aau93051 Arabidops
255	55	82.1	393	4	AAB70924	Aab70924 Polyoma v	328	55	82.1	798	5	AAU43116	Aau43116 A. thalia
256	55	82.1	419	8	ADR66870	Adr66870 Human pro	329	55	82.1	798	7	ADD30228	Add30228 Plant vie
257	55	82.1	419	8	ADR65972	Adr65972 Human pro	330	55	82.1	798	7	ADG37127	Adg37127 Plant vie
258	55	82.1	426	4	ABB11048	Abb11048 Human equ	331	55	82.1	798	8	ADI41767	Adi41767 Plant tra
259	55	82.1	434	3	AAB42352	Aab42352 Human ORF	332	55	82.1	798	8	ADI61347	Adi61347 A. thalia
260	55	82.1	451	7	ADE62342	Ade62342 Rat Prote	333	55	82.1	798	8	ADO02319	Ado02319 Thalecres
261	55	82.1	454	7	ABM87734	Abm87734 Rice abio	334	55	82.1	834	7	ADC31026	Adc31026 Human nov
262	55	82.1	456	6	ABP98836	Abp98836 Human str	335	55	82.1	844	7	ADR08920	Adr08920 Human pro
263	55	82.1	467	7	ADM04921	Adm04921 Human pro	336	55	82.1	848	5	ABG66688	Abg66688 Human nov
264	55	82.1	473	3	AAG24963	Aag24963 Arabidops	337	55	82.1	849	6	ADA13338	Ada13338 Human int
265	55	82.1	473	3	AAG43481	Aag43481 Arabidops	338	55	82.1	854	3	AAG42329	Aag42329 Arabidops
266	55	82.1	473	8	ADT55517	Adt55517 Plant pol	339	55	82.1	854	7	ADD46674	Add46674 Human pro
267	55	82.1	494	4	ABG29246	Abg29246 Novel hum	340	55	82.1	906	5	ABG66687	Abg66687 Human nov
268	55	82.1	498	7	ABM86962	Abm86962 Rice abio	341	55	82.1	907	6	ABG74229	Abg74229 Canine An
269	55	82.1	512	8	ADY13503	Ady13503 Plant ful	342	55	82.1	911	7	ADE56304	Ade56304 Human pro
270	55	82.1	515	8	ADS12177	Ads12177 Human the	343	55	82.1	983	2	AAU09513	Aau09513 Mouse JMY
271	55	82.1	527	4	AAB70917	Aab70917 L. monocy	344	55	82.1	986	7	ADD01164	Add01164 Human nuc
272	55	82.1	533	4	AAB95397	Aab95397 Human pro	345	55	82.1	996	9	AEA02082	Aea02082 Novel hum
273	55	82.1	541	2	AAW37148	Aaw37148 Mammalian	346	55	82.1	1001	6	ABU98812	Abu98812 Novel hum
274	55	82.1	542	8	ADX79745	Adx79745 Plant ful	347	55	82.1	1001	8	ADO09372	Ado09372 Novel hum
275	55	82.1	552	2	AAW90172	Aaw90172 Human hea	348	55	82.1	1005	2	AAW93955	Aaw93955 Human 53B
276	55	82.1	552	4	AAAB86278	Aaab86278 Human DCM	349	55	82.1	1005	6	ABU98810	Abu98810 Novel hum
277	55	82.1	552	5	AAE25975	Aae25975 Human pro	350	55	82.1	1005	6	ABG71757	Abg71757 Human can
278	55	82.1	552	7	ADHG2560	Adhg2560 Cardiac m	351	55	82.1	1005	6	ADA10593	Ada10593 Human can
279	55	82.1	552	9	AEA21095	Aea21095 Novel hum	352	55	82.1	1005	8	ADO09368	Ado09368 Novel hum
280	55	82.1	554	6	ADA48166	Ada48166 Rice prot	353	55	82.1	1005	9	ADX05602	Adx05602 Cyclin-de
281	55	82.1	559	4	AAB20494	Aab20494 Human Pab	354	55	82.1	1044	8	ADR09546	Adr09546 Human pro
282	55	82.1	559	4	AAB73399	Aab73399 Human WAV	355	55	82.1	1051	3	AAG42328	Aag42328 Arabidops
283	55	82.1	559	4	AAG67345	Aag67345 Amino aci	356	55	82.1	1051	8	ADT55608	Adt55608 Plant pol
284	55	82.1	559	4	AAM52317	Aam52317 Human Sca	357	55	82.1	1057	3	AAG42327	Aag42327 Arabidops
285	55	82.1	559	7	ADD08985	Ad08985 Human pab	358	55	82.1	1068	6	ABU98811	Abu98811 Novel hum
286	55	82.1	559	7	ADJ68719	Adj68719 Human hea	359	55	82.1	1068	8	ADO09370	Ado09370 Novel hum
287	55	82.1	559	7	ADJ70867	Adj70867 Human hea	360	55	82.1	1070	4	ABG02155	Abg02155 Novel hum
288	55	82.1	559	8	ADG65216	Adg65216 Mouse Pab	361	55	82.1	1082	7	ADE14368	Ade14368 Human int
289	55	82.1	559	8	ADG65208	Adg65208 Human Pab	362	55	82.1	1088	8	ADH61284	Adh61284 INTSIG pr
290	55	82.1	559	8	ADR16249	Adr16249 Human Pab	363	55	82.1	1101	8	ADR66324	Adr66324 Human pro
291	55	82.1	559	8	ADS34608	Ads34608 POSH prot	364	55	82.1	1101	8	ADR66666	Adr66666 Human pro
292	55	82.1	568	6	ADA54994	Ada54994 Human pro	365	55	82.1	1112	8	ADR66113	Adr66113 Human pro
293	55	82.1	575	8	ADT07559	Adt07559 Human col	366	55	82.1	1112	8	ADR66455	Adr66455 Human pro
294	55	82.1	577	8	ADT07561	Adt07561 Human col	367	55	82.1	1112	9	ADY18523	Ady18523 PRO poly
295	55	82.1	585	8	ADT07557	Adt07557 Human mod	368	55	82.1	1128	8	ADN12357	Adn12357 ASPP2 pol
296	55	82.1	601	7	ADB64412	Adb64412 Human pro	369	55	82.1	1128	9	AEA32205	Aea32205 Human ASP
297	55	82.1	619	4	ABE68682	Abbe68682 Drosophil	370	55	82.1	1128	9	AEI16419	Aei16419 ASPP2. 8/
298	55	82.1	629	4	AAU72168	Aau72168 Human RNA	371	55	82.1	1141	6	ABR39812	Ab39812 Human SCA
299	55	82.1	634	7	ADM04327	Adm04327 Human pro	372	55	82.1	1143	7	ADC31787	Adc31787 Human nov
300	55	82.1	638	8	ADI82510	Adi82510 Human mod	373	55	82.1	1152	9	ADY16987	Ade16987 Full leng
301	55	82.1	656	9	ADV99949	Adv99949 Human PEM	374	55	82.1	1171	5	AAE16987	Aae16987 Mouse DRF
302	55	82.1	659	9	ADV99948	Adv99948 Human PEM	375	55	82.1	1193	8	ADP04178	Adp04178 Human col
303	55	82.1	659	9	ADV99945	Adv99945 Human PEM	376	55	82.1	1196	6	ABR39811	Ab39811 Human SCA
304	55	82.1	675	6	AAO26721	Aao26721 SR protei	377	55	82.1	1203	8	ADT07558	Adt07558 Human col
305	55	82.1	681	7	ADE59481	Ade59481 Rat Prote	378	55	82.1	1205	8	ADT07560	Adt07560 Human col
306	55	82.1	686	5	AAU80373	Aau80373 Human cel	379	55	82.1	1213	8	ADT07556	Adt07556 Human col
307	55	82.1	686	6	AAE34823	Aae34823 Protein #	380	55	82.1	1216	8	ADX97588	Adx97588 Pancreati
308	55	82.1	686	8	ADO19347	Ado19347 Human pro	381	55	82.1	1248	2	AAU13464	Aau13464 Human dia
309	55	82.1	686	9	ADU05467	Adu05467 Cyclin-de	382	55	82.1	1250	6	AAO26231	Aao26231 MDT rela
310	55	82.1	686	9	ADY15012	Ady15012 PRO poly	383	55	82.1	1250	8	ADM87156	Adm87156 Human pro
311	55	82.1	686	9	ABE34464	Abe34464 SR protei	384	55	82.1	1255	2	AAW52249	Aaw52249 Mouse Rho
312	55	82.1	687	8	ADY07407	Ady07407 Plant ful	385	55	82.1	1255	2	AAW76733	Aaw76733 Mouse mDi
313	55	82.1	688	2	AAU27053	Aau27053 Human pro	386	55	82.1	1315	2	AAW76734	Aaw76734 Human mDi
314	55	82.1	688	6	AAO26722	Aao26722 SR protei	387	55	82.1	1311	4	ABG28241	Abg28241 Novel hum
315	55	82.1	688	8	ADJ96645	Adj96645 Human SRP	388	55	82.1	1375	8	ADB43646	Adb43646 Bacterial
316	55	82.1	691	7	ADF76663	Adf76663 Novel hum	389	55	82.1	1410	8	ABM81569	Abm81569 Tumour-as

390	55	82.1	1467	5	AAU78283	Aau78283 Human apo	463	52	77.6	41	8	ADU09353	Adu09353 Synthetic
391	55	82.1	1527	8	AAU01184	Aau01184 Rat Gluta	464	52	77.6	44	4	ABG16321	Abg16321 Novel hum
392	55	82.1	1527	8	ADF45523	Adf45523 Rat KIAA0	465	52	77.6	46	4	ABG17583	Abg17583 Novel hum
393	55	82.1	1592	8	ADL35711	Adl35711 Human ang	466	52	77.6	47	5	ADRA1538	Adra1538 Human CD-
394	55	82.1	1592	8	ABM81997	Abm81997 Tumour-as	467	52	77.6	53	6	ABB98725	Abb98725 Human PRI
395	55	82.1	1616	8	ADP22958	Adp22958 PRO polyp	468	52	77.6	53	6	ABB98733	Abb98733 Murine PR
396	55	82.1	1652	7	ADE71273	Adew71273 Novel hum	469	52	77.6	54	4	ABG05172	Abg05172 Novel hum
397	55	82.1	1682	8	ADRI14131	Adri14131 Human NF-	470	52	77.6	69	8	ADR94153	Adr94153 Novel S.
398	55	82.1	1765	7	ADB80341	Adb80341 Human MDD	471	52	77.6	69	9	AEA58023	Aea58023 Streptoco
399	55	82.1	1822	5	ABPA43899	Abp43899 Nuclear p	472	52	77.6	78	6	ABB98736	Abb98736 Murine PR
400	55	82.1	1822	7	ADE59297	Adew59297 Human Pro	473	52	77.6	78	6	ABB98728	Abb98728 Human PRI
401	55	82.1	1822	7	ADE59293	Adew59293 Human Pro	474	52	77.6	81	5	ABP34010	Abp34010 Human ORF
402	55	82.1	2062	7	ADC31187	Adc31187 Human nov	475	52	77.6	88	4	AAO02124	Aao02124 Human pol
403	55	82.1	2068	8	ADO62867	Ado62867 Transcrip	476	52	77.6	88	6	ABB98739	Abb98739 Murine PR
404	55	82.1	2150	8	ADZ09836	Adz09836 Human bre	477	52	77.6	88	6	ABB98731	Abb98731 Human PRI
405	55	82.1	2783	2	AAR23963	Aar23963 AFP-1 (Al	478	52	77.6	103	3	ABB33096	Abb33096 Pinus rad
406	55	82.1	2783	2	AAR23962	Aar23962 AFP-1. 3/	479	52	77.6	104	3	ABB78538	Abb78538 Ser-Pro-p
407	55	82.1	2783	4	AAB82946	Aab82946 Human and	480	52	77.6	104	8	ADU09325	Adu09325 [GPPPP]n
408	55	82.1	2789	7	ADJ70422	Adj70422 Human hea	481	52	77.6	105	4	AAO00771	Aao00771 Human pol
409	55	82.1	2992	8	ADP30190	Adp30190 Human sec	482	52	77.6	113	3	AAAG11750	Aag11750 Arabidops
410	55	82.1	3065	8	ADP30259	Adp30259 Human sec	483	52	77.6	113	6	ABB98732	Abb98732 Human PRI
411	55	82.1	3394	7	ADJ68723	Adj68723 Human hea	484	52	77.6	113	6	ABB98740	Abb98740 Murine PR
412	55	82.1	3572	5	ABG95659	Abg95659 Human nuc	485	52	77.6	114	7	ADC33200	Adc33200 Human nov
413	55	82.1	4873	6	ABO14747	Abol14747 Novel hum	486	52	77.6	118	6	ABB98730	Abb98730 Human PRI
414	54	80.6	171	7	ADF08027	Adf08027 Bacterial	487	52	77.6	118	6	ABB98738	Abb98738 Murine PR
415	54	80.6	242	8	ADY22812	Ady22812 Plant ful	488	52	77.6	121	2	AAR29164	Aar29164 PRP3. 3/2
416	54	80.6	378	8	ADX72267	Adx72267 Plant ful	489	52	77.6	127	3	AAAG06240	Aag06240 Arabidops
417	54	80.6	406	6	ABP58350	Abp58350 Human cel	490	52	77.6	127	3	AAAG0268	Aag0268 Arabidops
418	54	80.6	485	6	ABP58349	Abp58349 Human cel	491	52	77.6	128	3	AAAG0267	Aag0267 Arabidops
419	54	80.6	485	7	ADC31623	Adc31623 Human nov	492	52	77.6	128	3	AAAG06239	Aag06239 Arabidops
420	54	80.6	485	7	ADM05364	Adm05364 Human pro	493	52	77.6	128	3	AAAG07010	Aag07010 Arabidops
421	54	80.6	591	4	ABG031618	Abg031618 Novel hum	494	52	77.6	131	6	ABJ25281	Abj25281 Mouse BAC
422	54	80.6	722	8	ADQ65219	Adq65219 Novel hum	495	52	77.6	132	4	AAO05753	Aao05753 Human pol
423	54	80.6	825	7	ADJ70479	Adj70479 Human hea	496	52	77.6	133	4	ABG09692	Abg09692 Novel hum
424	54	80.6	900	3	AAB42321	Aab42321 Human ORF	497	52	77.6	137	4	ABG12622	Abg12622 Novel hum
425	54	80.6	1005	8	ADQ65183	Adq65183 Novel hum	498	52	77.6	138	3	AAAG11749	Aag11749 Arabidops
426	54	80.6	1091	4	ABBS58383	Abbs58383 Drosophil	499	52	77.6	138	4	ABG22949	Abg22949 Novel hum
427	54	80.6	1217	4	ABG09876	Abg09876 Novel hum	500	52	77.6	138	7	ADC33086	Adc33086 Human nov
428	54	80.6	1638	7	ADC31236	Adc31236 Human nov	501	52	77.6	140	8	ADX73862	Adx73862 Plant ful
429	54	80.6	1822	8	ADN21439	Adn21439 Bacterial	502	52	77.6	144	3	AAAG0266	Aag0266 Arabidops
430	54	80.6	2397	8	ADU02443	Adu02443 Novel hum	503	52	77.6	144	3	AAAG07009	Aag07009 Arabidops
431	53	79.1	29	4	AAO04677	Aao04677 Human pol	504	52	77.6	144	3	AAAG06238	Aag06238 Arabidops
432	53	79.1	56	4	ABG12672	Abg12672 Novel hum	505	52	77.6	151	2	AAR29165	Aar29165 PRP3 (fro
433	53	79.1	100	4	ABG25413	Abg25413 Novel hum	506	52	77.6	152	7	ADB75611	Adb75611 Prostata
434	53	79.1	149	4	AAO05346	Aao05346 Human pol	507	52	77.6	153	1	ABP90380	Abp90380 Sequence
435	53	79.1	239	4	ABBS3164	Abbs3164 Drosophil	508	52	77.6	153	4	ABB67456	Abb67456 Amino aci
436	53	79.1	245	7	ADJ11494	Adj11494 Rice prot	509	52	77.6	153	6	ABB98723	Abb98723 Human PRI
437	53	79.1	245	7	ADJ11826	Adj11826 Rice prot	510	52	77.6	153	6	ABB98724	Abb98724 Murine PR
438	53	79.1	263	8	ADT57062	Adt57062 Plant pol	511	52	77.6	156	3	AAAG07008	Aag07008 Arabidops
439	53	79.1	271	9	ABM96440	Abm96440 M. xanthu	512	52	77.6	158	4	AAAB67457	Aab67457 Amino aci
440	53	79.1	363	4	AAAM25490	Aam25490 Human pro	513	52	77.6	161	6	ABU19480	Abu19480 Protein e
441	53	79.1	400	7	ADC29743	Adc29743 B. juncea	514	52	77.6	167	3	AAAG11748	Aag11748 Arabidops
442	53	79.1	493	4	ABG12673	Abg12673 Novel hum	515	52	77.6	175	8	ADN42048	Adn42048 Amino aci
443	53	79.1	511	8	ADR94269	Adr94269 Novel S.	516	52	77.6	178	8	ADN42050	Adn42050 Amino aci
444	53	79.1	511	9	AEA58139	Aea58139 Streptoco	517	52	77.6	184	5	ABO60210	Abog60210 Human gen
445	53	79.1	1258	3	AAAY80120	Aay80120 Human Shi	518	52	77.6	189	5	ABG76914	Abg76914 Rice WRKY
446	53	79.1	1258	4	ABAB98987	Abab98987 Human typ	519	52	77.6	198	4	ABG19714	Abg19714 Novel hum
447	53	79.1	1258	8	ADJ66611	Adj66611 Inositol	520	52	77.6	203	7	ADF08002	Adf08002 Bacterial
448	53	79.1	1258	8	ABM81289	Abm81289 Tumour-as	521	52	77.6	205	6	ABU20105	Abu20105 Protein e
449	53	79.1	1302	8	ADE28331	Adew28331 Human KPP	522	52	77.6	206	4	AAAM42138	Aam42138 Human pol
450	53	79.1	1316	4	ABG22997	Abg22997 Novel hum	523	52	77.6	211	7	ADD30995	Add30995 Plant yie
451	53	79.1	1321	4	ABG25416	Abg25416 Novel hum	524	52	77.6	211	8	ADI43759	Adi43759 Plant tra
452	52	77.6	10	2	AAAB93322	Aar93322 ABI prote	525	52	77.6	211	8	ADOO1627	Adoo1627 Thalecra
453	52	77.6	10	5	ABB78428	Abb78428 Hydroxypr	526	52	77.6	220	2	AAAY21431	Aay21431 Human hig
454	52	77.6	10	8	ADU09142	Adu09142 Ser-Hyp4	527	52	77.6	228	4	AAU05775	Aau05775 Rice inve
455	52	77.6	19	1	AAAP81947	Aap81947 19 amino	528	52	77.6	230	8	ADU09357	Adu09357 Synthetic
456	52	77.6	20	2	AAAB38960	Aaw38960 Peptide r	529	52	77.6	233	3	AAAB07859	Aab07859 Amino aci
457	52	77.6	20	8	ADU79115	Adu79115 Human pep	530	52	77.6	245	3	AAAB43357	Aab43357 Human ORF
458	52	77.6	29	4	AAO02160	Aao02160 Human pol	531	52	77.6	245	8	ADP55743	Adp55743 Human PRO
459	52	77.6	29	4	ABG08613	Abg08613 Novel hum	532	52	77.6	257	2	AAAY25630	Aay25630 Phleum sp
460	52	77.6	30	4	AAO02051	Aao02051 Human pol	533	52	77.6	257	7	ADC34880	Adc34880 Timothy g
461	52	77.6	33	3	AAAY54915	Aay54915 RGD-conta	534	52	77.6	264	4	AAAB64417	Aab64417 Amino aci
462	52	77.6	41	2	AAAY01285	Aay01285 Peptide e	535	52	77.6	264	6	ADA54869	Ada54869 Human pro

536	52	77.6	264	7	ADM04325	Adm04325 Human pro	609	52	77.6	666	2	AAW72911	Myobacte
537	52	77.6	270	4	AAU16480	Aau16480 Human nov	610	52	77.6	666	2	AAU21928	Amino aci
538	52	77.6	270	6	ABU55549	Abu55549 Human nov	611	52	77.6	666	4	ABU19846	Myobacte
539	52	77.6	270	9	ADSL1801	Adsl1801 Human the	612	52	77.6	666	5	ABU05989	M. tuberc
540	52	77.6	271	9	ADY18585	Ady18585 PRO poly	613	52	77.6	666	6	ABU37020	Protein e
541	52	77.6	272	6	ABO14766	Abol14766 Novel hum	614	52	77.6	668	5	ABB93879	Herbicida
542	52	77.6	275	9	ADM17101	Adm17101 Eucalyptu	615	52	77.6	668	9	ADY91562	Thale cre
543	52	77.6	285	2	AAV25633	Aav25633 Phileum sp	616	52	77.6	681	4	ABG07912	Novel hum
544	52	77.6	285	3	ADC34883	Adc34883 Timothy g	617	52	77.6	708	7	ADC19791	Human can
545	52	77.6	293	3	AAB43674	Aab43674 Human can	618	52	77.6	711	8	ADJ17971	Human PMW
546	52	77.6	297	8	ADR08753	Adr08753 Human pro	619	52	77.6	721	8	ADY22716	Plant ful
547	52	77.6	301	4	ABB68068	Abb68068 Drosophil	620	52	77.6	730	8	ADO62718	Transcrip
548	52	77.6	316	5	AAE22089	Aae22089 Human nov	621	52	77.6	781	8	ADO62717	Transcrip
549	52	77.6	319	8	ABO59236	Abos59236 Human gen	622	52	77.6	800	8	ABO58564	Human gen
550	52	77.6	323	2	AAW00923	Aaw00923 HaSNPV po	623	52	77.6	809	5	ABU03210	Human pho
551	52	77.6	325	7	ABM85700	Abm85700 Mouse pro	624	52	77.6	809	8	ADIL15651	Human pho
552	52	77.6	331	4	ABM70820	Abm70820 Drosophil	625	52	77.6	809	8	ADJ25986	Human pho
553	52	77.6	352	7	ADM04873	Adm04873 Human pro	626	52	77.6	809	8	ADN97887	Human pho
554	52	77.6	353	5	AAE22090	Aae22090 Human nov	627	52	77.6	809	8	ADO50282	Human pho
555	52	77.6	354	5	ABP41838	Abp41838 Human ova	628	52	77.6	809	8	ADO34034	Human pho
556	52	77.6	356	4	AAW70701	Aaw70701 Human TAA	629	52	77.6	809	8	ADQ78350	Human pho
557	52	77.6	357	5	AAU85555	Aau85555 Clone #50	630	52	77.6	809	9	ABE85186	Human pho
558	52	77.6	357	5	AAU69431	Aau69431 Lung smal	631	52	77.6	828	4	ABB58368	Drosophil
559	52	77.6	357	6	ABU69527	Abu69527 Human lun	632	52	77.6	847	8	ADR58600	Human int
560	52	77.6	357	6	ABU66430	Abu66430 Lung canc	633	52	77.6	863	4	AAAB86463	Murine HC
561	52	77.6	357	7	ADH47336	Adh47336 Human lun	634	52	77.6	885	7	ADE08478	Novel pro
562	52	77.6	357	8	ADJ21255	Adj21255 Human lun	635	52	77.6	913	9	ABE119955	Novel hum
563	52	77.6	360	8	ADY09822	Ady09822 Plant ful	636	52	77.6	1013	8	ADP24491	PRO poly
564	52	77.6	390	3	ABAS8287	Abas8287 Lung canc	637	52	77.6	1017	4	AAAM40352	Human pol
565	52	77.6	391	7	ABM87129	Abm87129 Rice abio	638	52	77.6	1039	7	ADJ70453	Human hea
566	52	77.6	402	9	ADX83401	Adx83401 Human TEG	639	52	77.6	1039	7	ABM90450	Rice abio
567	52	77.6	403	8	ADX74016	Adx74016 Plant ful	640	52	77.6	1054	5	ABB79311	Human ova
568	52	77.6	418	4	AAW20496	Aaw20496 Human Pab	641	52	77.6	1066	3	AAU84877	Amino aci
569	52	77.6	418	6	ABU11825	Abu11825 Human MDD	642	52	77.6	1076	5	ABP69070	Human pol
570	52	77.6	418	8	ADG65212	Adg65212 Human Pab	643	52	77.6	1098	7	ADE60625	Rat Prote
571	52	77.6	421	6	AAO16115	Aao16115 Human can	644	52	77.6	1100	7	ADE54782	Human Pro
572	52	77.6	428	3	AAW23606	Aaw23606 Arabidops	645	52	77.6	1100	7	ADN04034	Antipsori
573	52	77.6	431	3	AAW23234	Aaw23234 Arabidops	646	52	77.6	1126	6	ABU70708	Human adi
574	52	77.6	431	3	AAW09613	Aaw09613 Arabidops	647	52	77.6	1126	8	ABM80766	Tumour-as
575	52	77.6	431	3	AAW143535	Adi143535 Plant tra	648	52	77.6	1126	8	ABM80767	Tumour-as
576	52	77.6	431	9	AEA26765	Aea26765 Stress to	649	52	77.6	1132	7	ADD71144	Human int
577	52	77.6	431	9	AEA26765	Aea26765 Stress to	650	52	77.6	1132	7	ADD71144	Human int
578	52	77.6	435	3	AAW09612	Aaw09612 Arabidops	651	52	77.6	1132	7	ADE60627	Human Pro
579	52	77.6	439	8	ADY25088	Ady25088 Plant ful	652	52	77.6	1132	7	ADE60623	Human Pro
580	52	77.6	440	3	AAW49463	Aaw49463 Arabidops	653	52	77.6	1132	8	ABM80765	Tumour-as
581	52	77.6	443	3	AAW49462	Aaw49462 Arabidops	654	52	77.6	1139	5	ABG97360	Human CGD
582	52	77.6	451	3	AAW35712	Aaw35712 Arabidops	655	52	77.6	1161	5	ABG97354	Human CGD
583	52	77.6	451	8	ADT56622	Adt56622 Plant pol	656	52	77.6	1179	8	ADSL10655	Human the
584	52	77.6	455	5	ABP73562	Abp73562 Candida a	657	52	77.6	1200	6	ADA55114	Human pro
585	52	77.6	470	4	ABG21932	Abg21932 Novel hum	658	52	77.6	1222	3	AAW36453	Arabidops
586	52	77.6	480	4	ABG12623	Abg12623 Novel hum	659	52	77.6	1239	4	ABG09877	Novel hum
587	52	77.6	483	8	ADU02754	Adu02754 Novel hum	660	52	77.6	1257	3	AAW36452	Arabidops
588	52	77.6	487	7	ADU19197	Adu19197 Human her	661	52	77.6	1275	3	AAW36451	Arabidops
589	52	77.6	488	3	AAW22819	Aaw22819 Arabidops	662	52	77.6	1312	5	ABP69782	Human pol
590	52	77.6	504	5	ABP41965	Abp41965 Human ova	663	52	77.6	1337	8	ABM82145	Tumour-as
591	52	77.6	520	4	AAW49336	Aaw49336 Murine WA	664	52	77.6	1343	7	ADJ69279	Human hea
592	52	77.6	520	4	AAW67350	Aaw67350 Amino aci	665	52	77.6	1362	8	ADJ66579	Bromodoma
593	52	77.6	520	4	AAW52318	Aaw52318 Murine WA	666	52	77.6	1362	9	ADV85499	Human bro
594	52	77.6	521	7	ABM85832	Abm85832 Mouse pro	667	52	77.6	1362	9	ADX07379	Cyclin-de
595	52	77.6	523	7	ADC81177	Adc81177 Human GPC	668	52	77.6	1386	7	ADE59291	Rat Prote
596	52	77.6	530	6	ADA55193	Ada55193 Human pro	669	52	77.6	1386	7	ADE59295	Rat Prote
597	52	77.6	551	8	ADQ65217	Adq65217 Novel hum	670	52	77.6	1390	5	ABU65152	Human nov
598	52	77.6	560	4	AAW95239	Aaw95239 Human pro	671	52	77.6	1390	8	ADN61955	Human nov
599	52	77.6	594	4	ABW61362	Abw61362 Drosophil	672	52	77.6	1400	8	ADV85501	Mouse bro
600	52	77.6	599	7	ADM03784	Adm03784 Human pro	673	52	77.6	1542	8	ADL83092	Human PRO
601	52	77.6	603	5	AAE22088	Aae22088 Human nov	674	52	77.6	1542	8	ADN04196	Antipsori
602	52	77.6	603	8	ADJ65163	Adj65163 Potassium	675	52	77.6	1542	8	ADP56292	Human PRO
603	52	77.6	620	9	ADY54945	Ady54945 Chronic v	676	52	77.6	1555	4	ABBS8847	PRO poly
604	52	77.6	622	8	ADJ65161	Adj65161 Potassium	677	52	77.6	1555	4	ABBS8847	Drosophil
605	52	77.6	649	7	ADC87385	Adc87385 Human GPC	678	52	77.6	1596	7	AAE37929	Human CGD
606	52	77.6	649	7	ADD46407	Add46407 Human pro	679	52	77.6	1966	7	ADJ70411	Human hea
607	52	77.6	651	9	ADY17208	Ady17208 PRO poly	680	52	77.6	2343	3	ABJ12453	Human HNR
608	52	77.6	656	5	AAE19183	Aae19183 Human pro	681	52	77.6	2440	8	ADQ89792	Antagonis

682	52	77.6	2440	9	ADW28417	Adw28417 Human nuc	755	51	76.1	242	4	AAM28950	Aam28950 Peptide #
683	52	77.6	2453	3	AAB12454	Aab12454 HNRCR pro	756	51	76.1	242	4	ABB30277	Abb30277 Peptide #
684	52	77.6	3119	2	AAW36888	Aaw36888 Mouse hun	757	51	76.1	242	4	ABB20882	Abb20882 Peptide #
685	52	77.6	3119	2	AAW44743	Aaw44743 Mouse hun	758	51	76.1	242	4	AAW68643	Aam68643 Human bon
686	52	77.6	3139	2	AAW08898	Aay08898 Human hun	759	51	76.1	242	4	AAW56266	Aam56266 Human bra
687	52	77.6	3232	4	ABB69873	Abb69873 Drosophil	760	51	76.1	242	4	ABG50303	Abg50303 Human liv
688	51	76.1	10	2	AAW93320	Aar93320 ABL prote	761	51	76.1	242	4	AAW04184	Aam04184 Peptide #
689	51	76.1	10	2	AAW93316	Aar93316 ABL prote	762	51	76.1	242	5	ABG38222	Abg38222 Human pep
690	51	76.1	29	1	AAW10194	Aap10194 Sequence	763	51	76.1	256	5	ABG33384	Abg33384 Human WT1
691	51	76.1	29	1	AAW10194	Aap10194 Sequence	764	51	76.1	256	7	ADB67537	Adb67537 Human wil
692	51	76.1	29	1	AAW80038	Aap80038 Beta-huma	765	51	76.1	256	7	ADJ80770	Adj80770 Wilm's tu
693	51	76.1	29	1	AAW91844	Aap91844 Analogue	766	51	76.1	256	8	ADJ83690	Adj83690 Human WT1
694	51	76.1	29	2	AAW42211	Aaw42211 Peptide d	767	51	76.1	256	8	ADL57578	Adl57578 Human WT-
695	51	76.1	29	2	AAW69457	Aaw69457 HCG anti	768	51	76.1	256	8	ADO09133	Ado09133 Human WT1
696	51	76.1	29	2	AAW93442	Aaw93442 Human hCG	769	51	76.1	256	8	ADO09133	Ado09133 Human WT1
697	51	76.1	29	3	AAW87487	Aay87487 Human cho	770	51	76.1	256	7	ABM86694	Abm86694 Rice abio
698	51	76.1	29	3	AAW20551	Aab20551 Human cho	771	51	76.1	272	8	ADX90852	Adx90852 Plant ful
699	51	76.1	29	4	AAU01147	Aau01147 Structure	772	51	76.1	278	2	AAW79095	Aar79095 Rat Fas l
700	51	76.1	29	4	AAW48393	Aab48393 Human cho	773	51	76.1	278	2	AAW98069	Aaw98069 Rat Fas l
701	51	76.1	29	4	AAU02845	Aau02845 Human cho	774	51	76.1	278	2	AAW95040	Aaw95040 Rat FasL
702	51	76.1	29	4	AAW04129	Aab04129 Peptide f	775	51	76.1	278	5	AAO19029	Aao19029 Human Fas
703	51	76.1	33	4	AAO04814	Aao04814 Human pol	776	51	76.1	278	7	ADD69944	Adp69944 Rat Fas l
704	51	76.1	41	4	AAO02061	Aao02061 Human pol	777	51	76.1	278	8	ADP26982	Adp26982 Rat FasL
705	51	76.1	44	4	AAO04929	Aao04929 Human pol	778	51	76.1	278	8	ADP26982	Adp26982 Rat FasL
706	51	76.1	69	7	ADC32722	Adc32722 Human nov	779	51	76.1	280	7	ADJ80896	Adj80896 Norway ra
707	51	76.1	70	3	AAW10803	Aag10803 Arabidops	780	51	76.1	280	8	ADJ83816	Adj83816 Wilm's tu
708	51	76.1	76	3	AAW10802	Aag10802 Arabidops	781	51	76.1	280	8	ADL57704	Adl57704 Human WT1
709	51	76.1	77	4	AAW92774	Aam92774 Human dig	782	51	76.1	280	8	ADO09259	Ado09259 Human WT-
710	51	76.1	80	4	AAO04412	Aao04412 Human pol	783	51	76.1	281	7	ADJ80900	Adj80900 Wilm's tu
711	51	76.1	83	8	ADI45175	Adi45175 Rice isop	784	51	76.1	281	8	ADJ83820	Adj83820 Recombina
712	51	76.1	93	4	ABG04608	Abg04608 Novel hum	785	51	76.1	281	8	ADL57708	Adl57708 Human WT-
713	51	76.1	99	4	AAO02076	Aao02076 Human pol	786	51	76.1	281	8	ADO09263	Ado09263 Human WT-
714	51	76.1	99	5	ABG33386	Abg33386 Human WT1	787	51	76.1	281	9	ABE55246	Aeb55246 Human b-z
715	51	76.1	99	7	ADB67544	Adb67544 Human pro	788	51	76.1	288	4	AAW80259	Aam80259 Human pro
716	51	76.1	99	7	ADJ80777	Adj80777 Wilm's tu	789	51	76.1	290	4	ABB68359	Abb68359 Drosophil
717	51	76.1	99	8	ADJ83697	Adj83697 Human WT1	790	51	76.1	291	7	ADJ80899	Adj80899 Wilm's tu
718	51	76.1	99	8	ADL57585	Adl57585 Human wil	791	51	76.1	291	8	ADJ83819	Adj83819 Recombina
719	51	76.1	99	8	ADO09140	Ado09140 Human WT1	792	51	76.1	291	8	ADL57707	Adl57707 Human WT-
720	51	76.1	109	4	AAO02138	Aao02138 Human pol	793	51	76.1	291	8	ADO09262	Ado09262 Human WT-
721	51	76.1	112	4	AAO02189	Aao02189 Human pol	794	51	76.1	292	7	ADJ80934	Adj80934 Wilm's tu
722	51	76.1	126	8	ADL06078	Adl06078 M. catarr	795	51	76.1	292	7	ADJ80905	Adj80905 Wilm's tu
723	51	76.1	128	4	AAO03145	Aao03145 Human pol	796	51	76.1	292	8	ADL57713	Adl57713 Human TAT
724	51	76.1	129	8	ADX94859	Adx94859 Plant ful	797	51	76.1	292	8	ADL57742	Adl57742 Human TAT
725	51	76.1	168	4	ABG13945	Abg13945 Novel hum	798	51	76.1	292	8	ADO09268	Ado09268 pStumpy-W
726	51	76.1	171	4	AAW17160	Aam17160 Peptide #	799	51	76.1	292	8	ADO09297	Ado09297 pStumpy-W
727	51	76.1	171	4	AAW15263	Aam15263 Peptide #	800	51	76.1	308	8	ADO09297	Ado09297 Plant ful
728	51	76.1	171	4	ABB36162	Abb36162 Peptide #	801	51	76.1	309	8	ADY04216	Ady04216 Plant ful
729	51	76.1	171	4	ABB34254	Abb34254 Peptide #	802	51	76.1	311	3	AAW17007	Aag17007 Arabidops
730	51	76.1	171	4	AAW29651	Aam29651 Peptide #	803	51	76.1	315	8	ADT56439	Adt56439 Plant pol
731	51	76.1	171	4	AAW27728	Aam27728 Peptide #	804	51	76.1	320	9	ADW17149	Adw17149 Eucalyptu
732	51	76.1	171	4	ABB30969	Abb30969 Peptide #	805	51	76.1	321	7	ADJ80914	Adj80914 Wilm's tu
733	51	76.1	171	4	ABB29093	Abb29093 Peptide #	806	51	76.1	321	7	ADJ80890	Adj80890 Wilm's tu
734	51	76.1	171	4	ABB19690	Abb19690 Protein #	807	51	76.1	321	8	ADJ83810	Adj83810 Human WT1
735	51	76.1	171	4	ABB21546	Abb21546 Protein #	808	51	76.1	321	8	ADL57698	Adl57698 Human wil
736	51	76.1	171	4	AAW67437	Aam67437 Human bon	809	51	76.1	321	8	ADL57722	Adl57722 Human TAT
737	51	76.1	171	4	AAW69331	Aam69331 Human bon	810	51	76.1	321	8	ADO09253	Ado09253 Human WT1
738	51	76.1	171	4	AAW56944	Aam56944 Human bra	811	51	76.1	321	7	ADO09277	Ado09277 Codon opt
739	51	76.1	171	4	AAW55047	Aam55047 Human bra	812	51	76.1	327	7	ADC87451	Adc87451 Human GPC
740	51	76.1	171	4	ABG49093	Abg49093 Human liv	813	51	76.1	338	3	AAW17006	Aag17006 Arabidops
741	51	76.1	171	4	ABG51005	Abg51005 Human liv	814	51	76.1	347	3	AAW42695	Aab42695 Human ORF
742	51	76.1	171	4	AAW03008	Aam03008 Peptide #	815	51	76.1	347	7	ADG39809	Adg39809 Protein s
743	51	76.1	171	4	AAW04854	Aam04854 Peptide #	816	51	76.1	363	3	AAW07704	Aab07704 Protein e
744	51	76.1	171	5	ABG37060	Abg37060 Human pep	817	51	76.1	377	8	ADJ80833	Adj80833 Human pro
745	51	76.1	171	5	ABG38947	Abg38947 Human pep	818	51	76.1	378	6	ABU07370	Abu07370 Human nov
746	51	76.1	175	4	AAO12806	Aao12806 Human pol	819	51	76.1	378	7	ADG39772	Adg39772 Human nov
747	51	76.1	186	8	ADK48382	Adk48382 Streptoco	820	51	76.1	391	3	AAW93904	Aay93904 Amino aci
748	51	76.1	199	7	ADC32888	Adc32888 Human nov	821	51	76.1	391	5	ABG92292	Abg92292 Herbicida
749	51	76.1	200	8	ADU01099	Adu01099 Human pro	822	51	76.1	391	8	ADT56626	Adt56626 Plant pol
750	51	76.1	200	8	ADU15481	Adu15481 Novel hum	823	51	76.1	404	8	ADX66693	Adx66693 Plant ful
751	51	76.1	210	2	AAW85066	Aaw85066 Wilm's tu	824	51	76.1	406	4	ABG27250	Abg27250 Novel hum
752	51	76.1	210	2	AAW22883	Aaw22883 Wilm's tu	825	51	76.1	408	8	ADX67722	Adx67722 Plant ful
753	51	76.1	242	4	AAW16455	Aam16455 Peptide #	826	51	76.1	409	3	AAW07702	Aab07702 Protein e
754	51	76.1	242	4	ABB35443	Abb35443 Peptide #	827	51	76.1	410	5	ABG33382	Abg33382 Human Trx

828	51	76.1	410	7	ADB67535	Adb67535 Human WT1	901	51	76.1	469	8	ADL57697	Adl57697 Human Wil
829	51	76.1	410	7	ADJ80768	Adj80768 Wilm's tu	902	51	76.1	469	8	ADO09252	Ado09252 Human WT1
830	51	76.1	410	8	ADJ83688	Adj83688 Human WT1	903	51	76.1	495	5	ABG33401	Abg33401 Human WT1
831	51	76.1	410	8	ADL57576	Adl57576 Human Trx	904	51	76.1	495	7	ADB67611	Adb67611 Human WT1
832	51	76.1	410	8	ADO09131	Ado09131 Human Trx	905	51	76.1	495	7	ADJ80844	Adj80844 Wilm's tu
833	51	76.1	417	5	ABG33399	Abg33399 Human WT1	906	51	76.1	495	8	ADJ83764	Adj83764 Human WT-
834	51	76.1	417	7	ADB67609	Adb67609 Human Wil	907	51	76.1	495	8	ADL57652	Adl57652 Human LAM
835	51	76.1	417	7	ADJ80842	Adj80842 Wilm's tu	908	51	76.1	495	8	ADO09207	Ado09207 Human LAM
836	51	76.1	417	7	ADJ83762	Adj83762 Human WT1	909	51	76.1	498	4	AAAB73400	Aab73400 Human WAV
837	51	76.1	417	8	ADL57650	Adl57650 Human Wil	910	51	76.1	498	8	ABM80048	Abm80048 Tumour-as
838	51	76.1	417	8	ADO09205	Ado09205 Mutant hu	911	51	76.1	498	9	ABE56517	Abe56517 Radiochem
839	51	76.1	420	5	ABG33393	Abg33393 Human Ral	912	51	76.1	502	8	ADP12999	Adp12999 Protein e
840	51	76.1	420	7	ADB67595	Adb67595 Human Ral	913	51	76.1	504	5	ABG33402	Abg33402 Human WT1
841	51	76.1	420	7	ADJ80828	Adj80828 Wilm's tu	914	51	76.1	504	7	ADB67612	Adb67612 Human WT1
842	51	76.1	420	8	ADJ83748	Adj83748 Human WT-	915	51	76.1	504	7	ADJ80845	Adj80845 Wilm's tu
843	51	76.1	420	8	ADL57636	Adl57636 Human Ral	916	51	76.1	504	8	ADJ83765	Adj83765 Human WT-
844	51	76.1	420	8	ADO09191	Ado09191 Human Ral	917	51	76.1	504	8	ADL57653	Adl57653 Human ubi
845	51	76.1	429	2	AAR85065	Aar85065 Wilm's tu	918	51	76.1	504	8	ADO09208	Ado09208 Human ubi
846	51	76.1	429	5	ABG33400	Abg33400 Human WT1	919	51	76.1	509	8	ADl42004	Adl42004 Plant tra
847	51	76.1	429	7	ADB67610	Adb67610 Human Wil	920	51	76.1	524	7	ABM89472	Abm89472 Rice abio
848	51	76.1	429	7	ADJ80843	Adj80843 Wilm's tu	921	51	76.1	568	5	ABG33392	Abg33392 Human Ral
849	51	76.1	429	8	ADJ83763	Adj83763 Human WT1	922	51	76.1	568	7	ADB67594	Adb67594 Human WT1
850	51	76.1	429	8	ADL57651	Adl57651 Human Wil	923	51	76.1	568	7	ADJ80827	Adj80827 Wilm's tu
851	51	76.1	429	8	ADO09206	Ado09206 Human WT1	924	51	76.1	568	8	ADJ83747	Adj83747 Human WT-
852	51	76.1	438	7	ADl21633	Adl21633 Novel hum	925	51	76.1	568	8	ADL57635	Adl57635 Human Ral
853	51	76.1	440	7	ABM90225	Abm90225 Rice abio	926	51	76.1	568	8	ADO09190	Ado09190 Human WT1
854	51	76.1	440	2	AAW47176	Aaw47176 Wilm's tu	927	51	76.1	577	4	AAO22140	Aao22140 Ral2-WT1
855	51	76.1	449	3	AAAY98805	Aay98805 Mouse WT1	928	51	76.1	579	8	ADR66346	Adr66346 Human pro
856	51	76.1	449	3	AAAY98804	Aay98804 Human WT1	929	51	76.1	579	8	ADR66688	Adr66688 Human pro
857	51	76.1	449	3	AAAY80196	Aay80196 Mouse Wil	930	51	76.1	591	5	ABG33383	Abg33383 Human Trx
858	51	76.1	449	3	AAAY80197	Aay80197 Human Wil	931	51	76.1	591	7	ADB67536	Adb67536 Human WT1
859	51	76.1	449	3	AAAG62136	Aag62136 Human WT1	932	51	76.1	591	7	ADJ80769	Adj80769 Wilm's tu
860	51	76.1	449	4	AAAG62137	Aag62137 Mouse WT1	933	51	76.1	591	8	ADJ83689	Adj83689 Human WT1
861	51	76.1	449	4	AAAU68904	Aau68904 Mouse Wil	934	51	76.1	591	8	ADL57577	Adl57577 Human Trx
862	51	76.1	449	4	AAU68903	Aau68903 Human Wil	935	51	76.1	591	8	ADO09132	Ado09132 Human Trx
863	51	76.1	449	5	ABG31644	Abg31644 Mouse Wil	936	51	76.1	592	2	AAW82661	Aaw82661 A. thalia
864	51	76.1	449	5	ABG33374	Abg33374 Mouse WT1	937	51	76.1	598	4	ABG14000	Abg14000 Novel hum
865	51	76.1	449	5	ABG33373	Abg33373 Human WT1	938	51	76.1	598	7	ADF60423	Adf60423 Human con
866	51	76.1	449	5	ABG33396	Abg33396 Human WT1	939	51	76.1	607	7	ADE08437	Ade08437 Novel pro
867	51	76.1	449	5	AAAG78446	Aag78446 Predicted	940	51	76.1	610	5	AAU76338	Aau76338 A. thalia
868	51	76.1	449	5	AAAG78444	Aag78444 Murine WT	941	51	76.1	615	3	AAU46152	Aau46152 Arabidops
869	51	76.1	449	6	ABBY99880	Abby99880 Mouse Wil	942	51	76.1	634	7	ADM04655	Adm04655 Human pro
870	51	76.1	449	6	ABG71423	Abg71423 Human Wil	943	51	76.1	642	7	ADC94190	Adc94190 E. faeciu
871	51	76.1	449	6	ABG71421	Abg71421 Murine Wi	944	51	76.1	647	8	ADU01120	Adu01120 Human pro
872	51	76.1	449	6	ABR44357	AbR44357 Human pro	945	51	76.1	647	8	ADU15502	Adu15502 Novel hum
873	51	76.1	449	6	ABR38970	AbR38970 Human tum	946	51	76.1	652	8	ADX96479	Adx96479 Plant ful
874	51	76.1	449	6	ABR38969	AbR38969 Mouse tum	947	51	76.1	679	3	AAG46151	Aag46151 Arabidops
875	51	76.1	449	6	ABR44448	AbR44448 Wilm's tu	948	51	76.1	691	3	AAG46150	Aag46150 Arabidops
876	51	76.1	449	7	ADB67606	Adb67606 Human Wil	949	51	76.1	692	4	AAG62154	Aag62154 Human WT1
877	51	76.1	449	7	ADB67522	Adb67522 Mouse Wil	950	51	76.1	697	8	ADX97270	Adx97270 Plant ful
878	51	76.1	449	7	ADB67521	Adb67521 Human Wil	951	51	76.1	729	8	ADY07005	Ady07005 Plant ful
879	51	76.1	449	7	ADJ80755	Adj80755 Wilm's tu	952	51	76.1	814	5	ABB92685	Abb92685 Herbicida
880	51	76.1	449	7	ADJ80839	Adj80839 Wilm's tu	953	51	76.1	857	4	ABB69603	Abb69603 Drosophil
881	51	76.1	449	7	ADJ80754	Adj80754 Wilm's tu	954	51	76.1	862	8	ADM66190	Adm66190 Novel hum
882	51	76.1	449	8	AAO24365	Aao24365 HLA-A24-r	955	51	76.1	883	8	ADM87622	Adm87622 Human EST
883	51	76.1	449	8	ADJ83674	Adj83674 Human WT1	956	51	76.1	963	7	ABM89266	Abm89266 Rice abio
884	51	76.1	449	8	ADJ83759	Adj83759 Human WT1	957	51	76.1	1002	8	ADO01055	Ado01055 Mouse hom
885	51	76.1	449	8	ADJ83675	Adj83675 Murine WT	958	51	76.1	1006	8	ADU20559	Adu20559 A. thalia
886	51	76.1	449	8	ADL57647	Adl57647 Human Wil	959	51	76.1	1006	8	ADU20561	Adu20561 A. thalia
887	51	76.1	449	8	ADL57562	Adl57562 Human Wil	960	51	76.1	1086	4	ABB69356	Abb69356 Drosophil
888	51	76.1	449	8	ADL57563	Adl57563 Mouse Wil	961	51	76.1	1228	3	AAG30727	Aag30727 Arabidops
889	51	76.1	449	8	ADN00725	Adn00725 Human WT1	962	51	76.1	1230	3	AAG30726	Aag30726 Arabidops
890	51	76.1	449	8	ADO09117	Ado09117 Human WT1	963	51	76.1	1248	3	AAG30725	Aag30725 Arabidops
891	51	76.1	449	8	ADO09118	Ado09118 Mouse WT1	964	51	76.1	1413	5	AAE21729	Aae21729 Human PKI
892	51	76.1	449	8	ADO09202	Ado09202 Human WT1	965	51	76.1	1431	9	ADV97898	Adv97898 Murine pr
893	51	76.1	449	8	ADO43852	Ado43852 Amino aci	966	51	76.1	1460	7	ADC59465	Adc59465 Novel hum
894	51	76.1	449	8	ADQ79585	Adq79585 Human Wil	967	51	76.1	1460	8	ADU04632	Adu04632 Human KIA
895	51	76.1	449	9	ADW72806	Adw72806 Human WT1	968	51	76.1	1518	6	ABU18375	Abu18375 Breast sp
896	51	76.1	449	9	AEA15654	Aea15654 Human WT1	969	51	76.1	1737	8	ADO01057	Ado01057 Human hom
897	51	76.1	453	7	ADB64865	Adb64865 Human pro	970	51	76.1	1772	3	AAG52485	Aag52485 Arabidops
898	51	76.1	467	4	ABG27768	Abg27768 Novel hum	971	51	76.1	1774	5	ABG69800	Abg69800 Human REM
899	51	76.1	469	7	ADJ80889	Adj80889 Wilm's tu	972	51	76.1	1776	3	AAG52484	Aag52484 Arabidops
900	51	76.1	469	8	ADJ83809	Adj83809 Human WT1	973	51	76.1	1787	3	AAG52483	Aag52483 Arabidops

974 51 76.1 1931 4 ABB66948 Drosophil
 975 51 76.1 2778 7 ADC51660 Human MEG
 976 51 76.1 2789 7 ADG75733 Human pro
 977 51 76.1 2854 5 ADH48732 NOV7 prot
 978 51 76.1 3051 6 ABR58308 BCU0947 p
 979 51 76.1 3302 8 ADL28831 Rat L-NAM
 980 51 76.1 3586 7 ADK40970 Novel hum
 981 51 76.1 3586 8 ADL15689 Kinase 72
 982 50.5 75.4 66 4 ABG16322 Novel hum
 983 50.5 75.4 150 4 ABG08615 Novel hum
 984 50 74.6 10 4 AAU09969 Proline 1
 985 50 74.6 11 2 AAU09969
 986 50 74.6 13 2 AAU09981 FAM-(Pro)
 987 50 74.6 13 4 AAB30652 Trypsin m
 988 50 74.6 15 2 AAW39030 Peptide x
 989 50 74.6 15 4 AAB30653 Trypsin m
 990 50 74.6 16 6 ABR62393 Synthetic
 991 50 74.6 18 2 AAW38907 Peptide x
 992 50 74.6 18 2 AAW38974 Peptide x
 993 50 74.6 18 4 AAB30654 Trypsin m
 994 50 74.6 19 2 AAW39009 Peptide x
 995 50 74.6 19 4 AAB30655 Trypsin m
 996 50 74.6 20 2 AAW38962 Peptide x
 997 50 74.6 20 2 AAW38971 Peptide x
 998 50 74.6 20 6 ABP59404 Self-asse
 999 50 74.6 20 6 ABP59405 Self-asse
 1000 50 74.6 21 4 AAB30656 Trypsin m

ALIGNMENTS

RESULT 1
 AAR93314
 ID AAR93314 standard; peptide; 10 AA.
 XX AC AAR93314;
 XX DT 24-APR-1996 (first entry)
 XX DE ABL protein tyrosine kinase derived peptide #1.
 XX KW SH3 ligand; SH3 binding agent; biased phage library;
 KW recognition sequence; src SH3 domain; Paget's disease; restenosis;
 KW rheumatoid arthritis; gout; neutrophil oxidase; tyrosine kinase; p47;
 KW p67; complex; chronic myelogenous leukaemia; cancer.
 XX OS Synthetic.
 XX WO9524419-A1.
 XX PD 14-SEP-1995.
 XX PF 13-MAR-1995; 95WO-US0033208.
 XX PR 11-MAR-1994; 94US-00209835.
 PR 06-JAN-1995; 95US-00369832.
 PA (ARIA-) ARIAD PHARM INC.
 XX PI Rickles RJ, Brugge JS, Botfield MC, Zoller MJ;
 XX WPI; 1995-328231/42.
 XX PT Identification of peptide(s) binding specifically to SH3 domains - for
 PT use in inhibiting interactions mediated by SH3 domains in treatment of
 PT e.g. osteoporosis and cancer.
 XX PS Disclosure; Fig 1; 74pp; English.
 XX CC The sequences given in AAR93272-342 represent peptides which are SH3
 CC ligands/SH3 binding agents. They represent a biased phage library which
 CC comprises six random amino acids flanking the tetrapeptide -ppip which

CC was identified as a recognition sequence for the src SH3 domain. These
 CC sequences were identified using the method of the invention. The method
 CC comprises contacting the SH3 domain with a mixture of peptides under
 CC conditions permitting a ligand to bind to an SH3 domain to form a
 CC complex. Any unbound peptides are removed and the complexed peptide
 CC ligands are dissociated from the complexes. The selected peptides are
 CC enriched by re-contacting them with the SH3 domain and then candidates
 CC which bind to the SH3 domain are detected. The isolated SH3 binding
 CC peptides may be used in the diagnosis, prevention and treatment of
 CC conditions or diseases resulting from cellular processes mediated by an
 CC SH3-based interaction. Such diseases include Paget's disease. Other
 CC conditions treatable with these peptides include restenosis, rheumatoid
 CC arthritis, gout and other problems in which an SH3 of neutrophil oxidase
 CC p47 and p67 complex is implicated, etc
 XX SQ Sequence 10 AA;
 Query Match 100.0%; Score 67; DB 2; Length 10;
 Best Local Similarity 100.0%; Pred. NO. 0.48;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 PPPYPPPPIP 10
 DB 1 PPPYPPPPIP 10
 RESULT 2
 ADA08234
 ID ADA08234 standard; peptide; 10 AA.
 XX AC ADA08234;
 XX DT 06-NOV-2003 (first entry)
 XX DE Human Abl-SH3 related peptide SEQ ID NO:138.
 XX KW identification; interacting protein; plurality; protein display library;
 KW binding protein; detection; specific binding;
 KW target-assisted iterative screening; protein-protein interaction.
 XX OS Synthetic.
 OS Homo sapiens.
 XX WO2003029821-A1.
 XX PD 10-APR-2003.
 XX PF 01-OCT-2002; 2002WO-US031349.
 XX PR 01-OCT-2001; 2001US-0326566P.
 XX PA (BUCK-) BUCK INST.
 XX PI Kourakine A, Bredesen D;
 XX WPI; 2003-403031/38.
 XX PT Identifying interacting proteins from a protein display library having
 PT potential binding proteins comprising contacting target proteins with the
 PT library and detecting binding of potential binding proteins with the
 PT target proteins.
 XX PS Example 1; Page 45; 61pp; English.
 XX CC The present invention describes a method for identifying interacting
 CC proteins from a plurality of potentially-interacting proteins comprising
 CC contacting one or more target proteins with a protein display library
 CC having a plurality of potential binding proteins and detecting specific
 CC binding of potential binding proteins with the target proteins. The
 CC method comprises: (a) contacting one or more target proteins with a
 CC protein display library comprising a plurality of potential binding
 CC proteins for one or more target proteins; (b) selecting members of the
 CC protein display library that bind to the target proteins to provide a

CC preselected set of potential binding proteins; (c) separating the members
 CC of the preselected set of potential binding proteins from the bound
 CC target protein and immobilising the members on a solid support such that
 CC the members are spatially addressable; (d) contacting the members of the
 CC preselected set of potential binding proteins with the target proteins;
 CC and (e) detecting specific binding of members of the preselected set of
 CC potential binding proteins with the target proteins, where binding of a
 CC member of the preselected set with a target protein indicates that the
 CC member and the target protein are interacting proteins. Also described is
 CC a kit for carrying out the method described above, comprising a protein
 CC display library and instructional materials providing protocols for the
 CC above method. The target-assisted iterative screening method is useful
 CC for determining large numbers of interactions (e.g. protein-protein
 CC interactions) between members of a library and various targets. The
 CC present sequence represents a peptide, which is used in the
 CC exemplification of the present invention.

XX
 SQ Sequence 10 AA;
 Query Match 100.0%; Score 67; DB 6; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.48; Mismatches 0; Indels 0; Gaps 0;
 Matches 10; Conservative 0;

QY 1 PPPYPPPPIP 10
 ID AAR93371
 DB 1 PPPYPPPPIP 10

RESULT 3
 AAR93369
 ID AAR93369 standard; peptide; 12 AA.

XX AAR93369;

XX 24-APR-1996 (first entry)

DE ABL protein tyrosine kinase derived peptide #1.

XX SH3 ligand; SH3 binding agent; biased phage library;
 KW recognition sequence; src SH3 domain; Paget's disease; restenosis;
 KW rheumatoid arthritis; gout; neutrophil oxidase; tyrosine kinase; p47;
 KW p67; complex; chronic myelogenous leukaemia; cancer.

XX Synthetic.

XX WO9524419-A1.

XX 14-SEP-1995.

XX 13-MAR-1995; 95WO-US003208.

XX 11-MAR-1994; 94US-00209835.

XX 06-JAN-1995; 95US-00369832.

XX (ARIA-) ARIAD PHARM INC.

XX Rickles RJ, Brugge JS, Botfield MC, Zoller MJ;

XX WPI; 1995-328231/42.

PT Identification of peptide(s) binding specifically to SH3 domains - for
 PT use in inhibiting interactions mediated by SH3 domains in treatment of
 PT e.g. osteoporosis and cancer.

XX Disclosure; Fig 2; 74pp; English.

XX The sequences given in AAR93369-76 represent peptides which are SH3
 CC ligands/SH3 binding agents. They represent a biased phage library which
 CC comprises six random amino acids flanking the hexapeptide PPPYP- which
 CC was identified as a recognition sequence for the src SH3 domain. These
 CC sequences were identified using the method of the invention. The method
 CC comprises contacting the SH3 domain with a mixture of peptides under
 CC conditions permitting a ligand to bind to an SH3 domain to form a

CC complex. Any unbound peptides are removed and the complexed peptide
 CC ligands are dissociated from the complexes. The selected peptides are
 CC enriched by re-contacting them with the SH3 domain and then candidates
 CC which bind to the SH3 domain are detected. The isolated SH3 binding
 CC peptides may be used in the diagnosis, prevention and treatment of
 CC conditions or diseases resulting from cellular processes mediated by an
 CC SH3-based interaction. Such diseases include Paget's disease. Other
 CC conditions treatable with these peptides include restenosis, rheumatoid
 CC arthritis, gout and other problems in which an SH3 of neutrophil oxidase
 CC p47 and p67 complex is implicated, etc

XX SQ Sequence 12 AA;

Query Match 100.0%; Score 67; DB 2; Length 12;

Best Local Similarity 100.0%; Pred. No. 0.56; Mismatches 0; Indels 0; Gaps 0;

Matches 10; Conservative 0;

QY 1 PPPYPPPPIP 10

ID AAR93371 standard; peptide; 12 AA.

DB 1 PPPYPPPPIP 10

RESULT 4

AAR93371

ID AAR93371 standard; peptide; 12 AA.

XX AAR93371;

XX 24-APR-1996 (first entry)

DE ABL protein tyrosine kinase derived peptide #3.

XX SH3 ligand; SH3 binding agent; biased phage library;

KW recognition sequence; src SH3 domain; Paget's disease; restenosis;

KW rheumatoid arthritis; gout; neutrophil oxidase; tyrosine kinase; p47;

KW p67; complex; chronic myelogenous leukaemia; cancer.

XX Synthetic.

XX WO9524419-A1.

XX 14-SEP-1995.

XX 13-MAR-1995; 95WO-US003208.

XX 11-MAR-1994; 94US-00209835.

XX 06-JAN-1995; 95US-00369832.

XX (ARIA-) ARIAD PHARM INC.

XX Rickles RJ, Brugge JS, Botfield MC, Zoller MJ;

XX WPI; 1995-328231/42.

PT Identification of peptide(s) binding specifically to SH3 domains - for
 PT use in inhibiting interactions mediated by SH3 domains in treatment of
 PT e.g. osteoporosis and cancer.

XX Disclosure; Fig 2; 74pp; English.

XX The sequences given in AAR93369-76 represent peptides which are SH3
 CC ligands/SH3 binding agents. They represent a biased phage library which
 CC comprises six random amino acids flanking the hexapeptide PPPYP- which
 CC was identified as a recognition sequence for the src SH3 domain. These
 CC sequences were identified using the method of the invention. The method
 CC comprises contacting the SH3 domain with a mixture of peptides under
 CC conditions permitting a ligand to bind to an SH3 domain to form a
 CC complex. Any unbound peptides are removed and the complexed peptide
 CC ligands are dissociated from the complexes. The selected peptides are
 CC enriched by re-contacting them with the SH3 domain and then candidates
 CC which bind to the SH3 domain are detected. The isolated SH3 binding
 CC peptides may be used in the diagnosis, prevention and treatment of
 CC conditions or diseases resulting from cellular processes mediated by an

CC SH3-based interaction. Such diseases include Paget's disease. Other
CC conditions treatable with these peptides include restenosis, rheumatoid
CC arthritis, gout and other problems in which an SH3 of neutrophil oxidase
CC p47 and p67 complex is implicated, etc
XX
SQ Sequence 12 AA;

Query Match 100.0%; Score 67; DB 2; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.56;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPPPPPPIP 10
Db 1 PPPPPPPPIP 10

RESULT 5
AAB17256
ID AAB17256 standard; peptide; 12 AA.
XX
AC AAB17256;
XX
DT 31-OCT-2000 (first entry)
XX
DE SH3 antagonist peptide sequence SEQ ID NO:312.
XX
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
KW thrombosis; pharmaceutical.
XX
OS Synthetic.
XX
PN WO200024782-A2.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US025044.
XX
PR 23-OCT-1998; 98US-0105371P.
PR 22-OCT-1999; 99US-00428082.
XX
PA (AMGB-) AMGEN INC.
XX
PI Peige U, Liu C, Cheatham J, Boone TC;
PI WPI; 2000-350702/30.
XX
DR
XX
PT Novel composition of matter comprising an Fc domain and pharmacologically
PT active peptides, useful for treating cancer and autoimmune diseases.
XX
PS Claim 39; Page 304; 608pp; English.
XX
CC The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)-a-F1-(X2)-b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)-c-P1, -(L1)-c-P1-(L2)-d-P2, -(L1)-c-P1-
CC (L2)-d-P2-(L3)-e-P3, or -(L1)-c-P1-(L2)-d-P2-(L3)-e-P3-(L4)-f-P4 where P1, P2,
CC P3, and P4 = are each independently sequences of pharmacologically active
CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,
CC c, d, e, and f = are each independently 0 or 1, provided that at least 1
CC of a and b is 1. The composition can have cytostatic, antiasthmatic,
CC thrombolytic and immunosuppressive activities. DNAs, vectors and host
CC cells from the present invention can be used for producing pharmaceutical
CC compositions. The compositions are useful for treating cancer, asthma,
CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
CC a Fab domain) can provide a longer half-life or incorporate functions
CC such as Fc receptor binding, protein A binding, complement fixation, and
CC possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to
CC AAB18003 represent nucleotide and amino acid sequences used in the

CC exemplification of the present invention
XX
SQ Sequence 12 AA;

Query Match 100.0%; Score 67; DB 3; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.56;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPPPPPPIP 10
Db 1 PPPPPPPPIP 10

RESULT 6
ABB73249
ID ABB73249 standard; peptide; 12 AA.
XX
AC ABB73249;
XX
DT 05-APR-2002 (first entry)
XX
DE Src homology3 (SH3) antagonist peptide SEQ ID NO:312.
XX
KW Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;
KW TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;
KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
KW cytostatic; antirheumatic; antiarthritis; antidiabetic; ophthalmological;
KW antianemic; anorectic; antifertility; haemostatic; dermatological;
KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
KW sleep disorder; neurological degenerative disease; anaemia;
KW thrombocytopaenia; metastatic tumour; systemic lupus erythematosus;
KW Fanconi's syndrome.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200183525-A2.
XX
PD 08-NOV-2001.
XX
PF 02-MAY-2001; 2001WO-US014310.
XX
PR 03-MAY-2000; 2000US-00563286.
XX
PA (AMGB-) AMGEN INC.
XX
PI Peige U, Liu C, Cheatham JC, Boone TC, Gudas JM;
PI WPI; 2002-130313/17.
XX
DR
XX
PT Novel vehicle-peptide molecule or its multimers useful for treating
PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
PT diabetic retinopathy, obesity, sleep disorders and infertility.
XX
PS Claim 39; Page 55; 176pp; English.
XX
CC The present invention describes a vehicle-peptide molecule (I) or its
CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,
CC cytostatic, antirheumatic, antiarthritis, antidiabetic, ophthalmological,
CC antianemic, anorectic, antifertility, haemostatic, dermatological and
CC neuroprotective activities. (I) can be used as a therapeutic or
CC prophylactic agent as well as for screening purposes. (I) is useful for
CC diagnosing diseases characterised by dysfunction of their associated
CC protein of interest, for identifying normal or abnormal proteins of
CC interest, as a part of diagnostic kit to detect the presence of their
CC proteins of interest in a biological sample. Additionally, (I) is useful
CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,
CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders, cancer,
CC infertility, and neurological degenerative diseases. (I), comprising EPO-
CC mimetic compounds are useful for treating disorders characterised by low

red blood cell levels such as anaemia. The TPO-mimetic comprising compounds are useful for treating conditions that involve an existing megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic tumour which result in thrombocytopaenia, systemic lupus erythematosus, and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777 represent amino acid and nucleic acid sequences used in the exemplification of the present invention

CC Sequence 12 AA;

Query Match 100.0%; Score 67; DB 5; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.56;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPYPPPPPIP 10
Db 1 PPPYPPPPPIP 10

RESULT 7
ADJ73403
ID ADJ73403 standard; peptide; 12 AA.
AC ADJ73403;
DT 06-MAY-2004 (first entry)
DE SH3 antagonist peptide sequence SeqID 858.
KW mimetic; CDR mimetibody; gene therapy; transgenic; immune;
KW cardiovascular; infectious; malignant; neurologic disease; anaemia;
KW immunomodulator; cardiant; antimicrobial; cytostatic; neuroprotective;
KW SH3.
OS Synthetic.
PN WO2003084477-A2.
PD 16-OCT-2003.
PP 24-MAR-2003; 2003WO-US009139.
PR 29-MAR-2002; 2002US-0368791P.
PA (CENZ) CENTOCOR INC.
PI Heavner GA, Knight DM, Scallion BJ, Ghayeb J;
XX WPI; 2003-804237/75.
XX New CDR mimetibody comprising a portion of a heavy or light chain
XX variable region comprising human framework or ligand binding region,
XX useful for preparing a composition for treating e.g., immune,
XX cardiovascular or neurologic disease.
XX Disclosure; SEQ ID NO 858; 97pp; English.

This invention relates to novel mammalian CDR mimetibodies, specific portions or variants thereof. Specifically, it refers to an antibody fragment where a protein has been inserted into, or replaces a portion of, one or more CDR regions, such that each CDR mimetibody comprises at least one portion of a heavy chain or light chain variable region, which itself comprises at least one human framework region and at least one ligand binding region (LBR). The present invention describes human mimetibodies, including modified immunoglobulins and cleavage products that can be useful in gene therapy and the generation of transgenic plants and animals. Furthermore, the CDR mimetibody is useful for preparing compositions for modulating, treating or reducing the symptoms of immune, cardiovascular, infectious, malignant and/or neurologic diseases, as well as anaemia. Accordingly, they exhibit immunomodulator, cardiant, antimicrobial, cytostatic and neuroprotective activities. This peptide sequence is an SH3 antagonist peptide sequence used to make a

mimetibody of the invention.

CC Sequence 12 AA;

Query Match 100.0%; Score 67; DB 7; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.56;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPYPPPPPIP 10
Db 1 PPPYPPPPPIP 10

RESULT 8
ADJ53037
ID ADJ53037 standard; peptide; 12 AA.
XX AC ADJ53037;
XX DT 06-MAY-2004 (first entry)
XX CH1 deleted mimetibody-related peptide SeqID858.
DE CH1 deleted mimetibody; immunosuppressive; cardiovascular; cardiant;
XX hypotensive; neuroprotective; nootropic; antibacterial; virucide;
KW fungicide; gene therapy; immune disorder; cardiovascular disease;
KW arrhythmia; hypertension; heart failure; neurodegenerative;
KW multiple sclerosis; dementia; Alzheimer's disease; anaemia;
KW cancerous condition; infectious disease; bacterial infection;
KW viral infection; fungal infection.
OS Unidentified.
OS Synthetic.
FH Key Location/Qualifiers
FT Misc-difference 1..12 /note= "All Xaa's in this sequence are unidentified amino acids"
FT
FT
XX WO2004002417-A2.
XX 08-JAN-2004.
XX 27-JUN-2003; 2003WO-US020347.
XX 28-JUN-2002; 2002US-0392431P.
XX (CENZ) CENTOCOR INC.
XX Heavner GA, Knight DM, Ghayeb J, Scallion BJ, Nesspor TC;
XX Kutoloski KA;
XX WPI; 2004-082870/08.
XX New CH1-deleted mimetibody polypeptides and nucleic acids, useful for
XX modulating, treating, alleviating, preventing an immune, cardiovascular,
XX or neurodegenerative disease or disorder, anemia, cancer, or infectious
XX diseases.
XX Claim 3; SEQ ID NO 858; 129pp; English.

This invention relates to CH1 deleted mimetibodies (and the DNA sequences which encode them), compositions, methods and uses. The invention may be useful for the development of compounds with an immunosuppressive, cardiovascular, cardiant, hypotensive, neuroprotective, nootropic, antibacterial, virucide or fungicide activity. In addition, the disclosed sequences may prove useful for gene therapy. The CH1-deleted mimetibody is useful for diagnosing or treating a disease condition in a cell, tissue, organ or animal, specifically for modulating the symptoms of an alleviating, preventing the incidence or reducing the symptoms of an immune, cardiovascular (for example arrhythmia, hypertension or heart failure), or neurodegenerative (for example multiple sclerosis, dementia or Alzheimer's disease) diseases or disorders, anaemia, cancerous

CC conditions, or infectious diseases (for example bacterial, viral or
 CC fungal infection). The present sequence is that of a peptide which may be
 CC used during the creation of a mimetibody of the invention.

XX
 SQ Sequence 12 AA;

Query Match 100.0%; Score 67; DB 8; Length 12;
 Best Local Similarity 100.0%; Pred. No. 0.56;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPYPPPPIP 10
 |||||
 Db 1 PPPYPPPPIP 10

RESULT 9
 ADJ51998
 ID ADJ51998 standard; peptide; 12 AA.
 AC ADJ51998;
 XX
 XX
 DT 06-MAY-2004 (first entry)
 XX
 XX CHI deleted mimetibody-related peptide SeqID858.

DE
 XX CHI deleted mimetibody; osteopathic; cardiovascular-Gen;
 KW dermatological-Gen; auditory; endocrine-Gen; gastrointestinal-Gen;
 KW gynaecological-Gen; hepatotropic; haemostatic; immunomodulator;
 KW antiallergic; muscular-Gen; cytostatic; antiinflammatory; neuroleptic;
 KW ophthalmologic; nephrotropic; respiratory-Gen; tumour necrosis factor;
 KW TNF; cytokine; bone disorder; joint disorder; cardiovascular disorder;
 KW dental disorder; oral disorder; dermatological disorder; ear disorder;
 KW nose disorder; throat disorder; endocrine disorder; metabolic disorder;
 KW gastrointestinal disorder; gynaecological disorder; hepatic disorder;
 KW obstetric disorder; haematologic disorder; immunological disorder;
 KW allergic disorder; infectious disorder; musculoskeletal disorder;
 KW oncological disorder; neurological disorder; nutritional disorder;
 KW ophthalmologic disorder; pediatric disorder; psychiatric disorder;
 KW renal disorder; pulmonary disorder.

XX
 OS Unidentified.
 OS Synthetic.

XX
 FH Key Location/Qualifiers
 FT Misc-difference 1. 12
 FT /notes "All Xaa's in this sequence are unidentified amino
 FT acids"

XX
 XX WO2004002424-A2.
 XX
 XX 08-JAN-2004.
 XX
 XX 30-JUN-2003; 2003WO-US020495.
 XX
 XX 28-JUN-2002; 2002US-0392431P.
 XX
 XX 19-SEP-2002; 2002US-0412144P.
 XX
 XX (CENZ) CENTOCOR INC.
 XX
 XX Heavner GA, Knight DM, Ghraryeb J, Scallion BJ, Nesspor TC;
 PI Kutoloski KA;
 PI
 XX WPI; 2004-082872/08.
 XX
 XX New CHI deleted mimetibody polypeptide and nucleic acid, useful for
 FT diagnosing, preventing or treating cardiovascular, dermatologic,
 FT endocrine, gastrointestinal, gynecologic, infectious, neurologic and
 FT nutritional disorders.

XX
 XX Claim 15; SEQ ID NO 858; 123pp; English.
 PS
 XX This invention relates to CHI deleted mimetibodies (and the DNA sequences
 CC which encode them), compositions, methods and uses. The invention may be
 CC

CC useful for the development of compounds with an osteopathic,
 CC cardiovascular-Gen, dermatological-Gen, auditory, endocrine-Gen,
 CC gastrointestinal-Gen, gynaecological-Gen, hepatotropic, haemostatic,
 CC immunomodulator, antiallergic, muscular-Gen, cytostatic,
 CC antiinflammatory, neuroleptic, ophthalmologic, nephrotropic or
 CC respiratory-Gen activity acting as a tumour necrosis factor (TNF)-
 CC modulator or cytokine-agonist. The methods and compositions of the
 CC present invention are useful for the diagnosis, prevention and/or
 CC treatment of diseases or conditions associated with aberrant expression
 CC or activity of the CHI deleted mimetibody, such as a bone or joint,
 CC cardiovascular, dental or oral, dermatological, ear, nose or throat,
 CC endocrine, metabolic, gastrointestinal, gynaecological, hepatic,
 CC obstetric, haematologic, immunological, allergic, infectious,
 CC musculoskeletal, oncological, neurological, nutritional, ophthalmologic,
 CC pediatric, psychiatric, renal or pulmonary disorders. The present
 CC sequence is that of a peptide which may be used during the creation of a
 CC mimetibody of the invention.

XX
 SQ Sequence 12 AA;

Query Match 100.0%; Score 67; DB 8; Length 12;
 Best Local Similarity 100.0%; Pred. No. 0.56;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPYPPPPIP 10
 |||||
 Db 1 PPPYPPPPIP 10

RESULT 10
 AAR93373
 ID AAR93373 standard; peptide; 12 AA.
 AC AAR93373;
 XX
 XX 24-APR-1996 (first entry)
 DT
 XX
 DE ABL protein tyrosine kinase derived peptide #5.
 XX
 KW SH3 ligand; SH3 binding agent; biased phage library;
 KW recognition sequence; src SH3 domain; Paget's disease; restenosis;
 KW rheumatoid arthritis; gout; neutrophil oxidase; tyrosine kinase; p47;
 KW p67; complex; chronic myelogenous leukaemia; cancer.

XX
 OS Synthetic.
 XX
 XX WO9524419-A1.
 XX
 XX 14-SEP-1995.
 XX
 XX 13-MAR-1995; 95WO-US003208.
 XX
 XX 11-MAR-1994; 94US-00209835.
 XX
 XX 06-JAN-1995; 95US-00369832.
 XX
 XX (ARIA-) ARIAD PHARM INC.
 XX
 XX Rickles RJ, Brugge JS, Botfield MC, Zoller MJ;
 XX WPI; 1995-328231/42.
 XX
 XX Identification of peptide(s) binding specifically to SH3 domains - for
 PT use in inhibiting interactions mediated by SH3 domains in treatment of
 PT e.g. osteoporosis and cancer.
 PT
 XX Disclosure; Fig 2; 74pp; English.
 PS
 XX The sequences given in AAR93369-76 represent peptides which are SH3
 CC ligands/SH3 binding agents. They represent a biased phage library which
 CC comprises six random amino acids flanking the hexapeptide PPPYP- which
 CC was identified as a recognition sequence for the src SH3 domain. These
 CC sequences were identified using the method of the invention. The method
 CC comprises contacting the SH3 domain with a mixture of peptides under

CC conditions permitting a ligand to bind to an SH3 domain to form a
 CC complex. Any unbound peptides are removed and the complexed peptide
 CC ligands are dissociated from the complexes. The selected peptides are
 CC enriched by re-contacting them with the SH3 domain and then candidates
 CC which bind to the SH3 domain are detected. The isolated SH3 binding
 CC peptides may be used in the diagnosis, prevention and treatment of
 CC conditions or diseases resulting from cellular processes mediated by an
 CC SH3-based interaction. Such diseases include Paget's disease. Other
 CC conditions treatable with these peptides include restenosis, rheumatoid
 CC arthritis, gout and other problems in which an SH3 of neutrophil oxidase
 CC p47 and p67 complex is implicated, etc
 XX
 SQ Sequence 12 AA;

Query Match 98.5%; Score 66; DB 2; Length 12;
 Best Local Similarity 90.0%; Pred. No. 0.71;
 Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPVPPPPIP 10
 |||||:|
 Db 1 PPPVPPPPVP 10

RESULT 11
 AAR93372
 ID AAR93372 standard; peptide; 12 AA.

XX AC AAR93372;

XX DT 24-APR-1996 (first entry)

XX DE ABL protein tyrosine kinase derived peptide #4.

XX KW SH3 ligand; SH3 binding agent; biased phage library;
 KW recognition sequence; src SH3 domain; Paget's disease; restenosis;
 KW rheumatoid arthritis; gout; neutrophil oxidase; tyrosine kinase; p47;
 KW p67; complex; chronic myelogenous leukaemia; cancer.

XX OS Synthetic.

XX PN WO9524419-A1.

XX PD 14-SEP-1995.

XX PF 13-MAR-1995; 95WO-US003208.

XX PR 11-MAR-1994; 94US-00209835.

XX PR 06-JAN-1995; 95US-00369832.

XX PA (ARIA-) ARIAD PHARM INC.

XX PI Rickles RJ, Brugge JS, Botfield MC, Zoller MJ;

XX DR WPI; 1995-328231/42.

XX PT Identification of peptide(s) binding specifically to SH3 domains - for
 XX use in inhibiting interactions mediated by SH3 domains in treatment of
 XX e.g. osteoporosis and cancer.

XX PS Disclosure; Fig 2; 74pp; English.

XX CC The sequences given in AAR93369-76 represent peptides which are SH3
 CC ligands/SH3 binding agents. They represent a biased phage library which
 CC comprises six random amino acids flanking the hexapeptide PPPVP- which
 CC was identified as a recognition sequence for the src SH3 domain. These
 CC sequences were identified using the method of the invention. The method
 CC comprises contacting the SH3 domain with a mixture of peptides under
 CC conditions permitting a ligand to bind to an SH3 domain to form a
 CC complex. Any unbound peptides are removed and the complexed peptide
 CC ligands are dissociated from the complexes. The selected peptides are
 CC enriched by re-contacting them with the SH3 domain and then candidates
 CC which bind to the SH3 domain are detected. The isolated SH3 binding
 CC peptides may be used in the diagnosis, prevention and treatment of

CC conditions or diseases resulting from cellular processes mediated by an
 CC SH3-based interaction. Such diseases include Paget's disease. Other
 CC conditions treatable with these peptides include restenosis, rheumatoid
 CC arthritis, gout and other problems in which an SH3 of neutrophil oxidase
 CC p47 and p67 complex is implicated, etc
 XX
 SQ Sequence 12 AA;

Query Match 98.5%; Score 66; DB 2; Length 12;
 Best Local Similarity 90.0%; Pred. No. 0.71;
 Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPVPPPPIP 10
 |||||:|
 Db 1 PPPVPPPPVP 10

RESULT 12
 AAR93370
 ID AAR93370 standard; peptide; 12 AA.

XX AC AAR93370;

XX DT 24-APR-1996 (first entry)

XX DE ABL protein tyrosine kinase derived peptide #2.

XX KW SH3 ligand; SH3 binding agent; biased phage library;
 KW recognition sequence; src SH3 domain; Paget's disease; restenosis;
 KW rheumatoid arthritis; gout; neutrophil oxidase; tyrosine kinase; p47;
 KW p67; complex; chronic myelogenous leukaemia; cancer.

XX OS Synthetic.

XX PN WO9524419-A1.

XX PD 14-SEP-1995.

XX PF 13-MAR-1995; 95WO-US003208.

XX PR 11-MAR-1994; 94US-00209835.

XX PR 06-JAN-1995; 95US-00369832.

XX PA (ARIA-) ARIAD PHARM INC.

XX PI Rickles RJ, Brugge JS, Botfield MC, Zoller MJ;

XX DR WPI; 1995-328231/42.

XX PT Identification of peptide(s) binding specifically to SH3 domains - for
 XX use in inhibiting interactions mediated by SH3 domains in treatment of
 XX e.g. osteoporosis and cancer.

XX PS Disclosure; Fig 2; 74pp; English.

XX CC The sequences given in AAR93369-76 represent peptides which are SH3
 CC ligands/SH3 binding agents. They represent a biased phage library which
 CC comprises six random amino acids flanking the hexapeptide PPPVP- which
 CC was identified as a recognition sequence for the src SH3 domain. These
 CC sequences were identified using the method of the invention. The method
 CC comprises contacting the SH3 domain with a mixture of peptides under
 CC conditions permitting a ligand to bind to an SH3 domain to form a
 CC complex. Any unbound peptides are removed and the complexed peptide
 CC ligands are dissociated from the complexes. The selected peptides are
 CC enriched by re-contacting them with the SH3 domain and then candidates
 CC which bind to the SH3 domain are detected. The isolated SH3 binding
 CC peptides may be used in the diagnosis, prevention and treatment of
 CC conditions or diseases resulting from cellular processes mediated by an
 CC SH3-based interaction. Such diseases include restenosis, rheumatoid
 CC conditions treatable with these peptides include restenosis, rheumatoid
 CC arthritis, gout and other problems in which an SH3 of neutrophil oxidase
 CC p47 and p67 complex is implicated, etc

XX

```
SQ Sequence 12 AA;
Query Match 98.5%; Score 66; DB 2; Length 12;
Best Local Similarity 90.0%; Pred. No. 0.71;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPYPPPPPIP 10
Db 1 PPPYPPPPVP 10
    |||||:|

RESULT 13
AAB17257
ID AAB17257 standard; peptide; 12 AA.
XX
AC AAB17257;
XX
XX 31-OCT-2000 (first entry)
XX
XX SH3 antagonist peptide sequence SEQ ID NO:313.
XX
XX Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
XX autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
XX immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
XX inhibitor; erythropoietin; thrombopoietin; interleukin 1;
XX cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
XX vascular endothelial growth factor; matrix metalloproteinase; asthma;
XX thrombosis; pharmaceutical.
XX
OS Synthetic.
XX
XX WO200024782-A2.
XX
XX 04-MAY-2000.
XX
XX 25-OCT-1999; 99WO-US025044.
XX
XX 23-OCT-1998; 98US-0105371P.
XX
XX 22-OCT-1999; 99US-00428082.
XX
XX (AMGE-) AMGEN INC.
XX
XX Feige U, Liu C, Cheatham J, Boone TC;
XX WPI; 2000-350702/30.
XX
XX Novel composition of matter comprising an Fc domain and pharmacologically
XX active peptides, useful for treating cancer and autoimmune diseases.
XX
XX Claim 39; Page 304; 608pp; English.
XX
XX The present invention describes composition of matter (I) comprising an
XX Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
XX (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
XX independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
XX (L2)d-P2-(L3)e-P3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
XX P3, and P4 = are each independently sequences of pharmacologically active
XX peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,
XX c, d, e, and f = are each independently 0 or 1, provided that at least 1
XX of a and b is 1. The composition can have cytostatic, antiasthmatic,
XX thrombolytic and immunosuppressive activities. DNAs, vectors and host
XX cells from the present invention can be used for producing pharmaceutical
XX compositions. The compositions are useful for treating cancer, asthma,
XX thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
XX a Fab domain) can provide a longer half-life or incorporate functions
XX such as Fc receptor binding, protein A binding, complement fixation, and
XX possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to
XX AAB18003 represent nucleotide and amino acid sequences used in the
XX exemplification of the present invention
XX
XX Sequence 12 AA;
Query Match 98.5%; Score 66; DB 3; Length 12;

Best Local Similarity 90.0%; Pred. No. 0.71;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPYPPPPPIP 10
Db 1 PPPYPPPPVP 10
    |||||:|

RESULT 14
ABB73250
ID ABB73250 standard; peptide; 12 AA.
XX
AC ABB73250;
XX
XX 05-APR-2002 (first entry)
XX
XX Src homology3 (SH3) antagonist peptide SEQ ID NO:313.
XX
XX Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
XX erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
XX TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;
XX TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;
XX MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
XX cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
XX antianemic; anorectic; antitumor; antitumor; haemostatic; dermatological;
XX neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
XX cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
XX sleep disorder; neurological degenerative disease; anaemia;
XX thrombocytopaenia; metastatic tumour; systemic lupus erythematosus;
XX Fanconi's syndrome.
XX
OS Homo sapiens.
XX
XX Synthetic.
XX
XX WO200183525-A2.
XX
XX 08-NOV-2001.
XX
XX 02-MAY-2001; 2001WO-US014310.
XX
XX 03-MAY-2000; 2000US-00563286.
XX
XX (AMGE-) AMGEN INC.
XX
XX Feige U, Liu C, Cheatham JC, Boone TC, Gudas JM;
XX WPI; 2002-130313/17.
XX
XX Novel vehicle-peptide molecule or its multimers useful for treating
XX inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
XX diabetic retinopathy, obesity, sleep disorders and infertility.
XX
XX Claim 39; Page 55; 176pp; English.
XX
XX The present invention describes a vehicle-peptide molecule (I) or its
XX multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,
XX cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,
XX antianemic, anorectic, antinfertility, haemostatic, dermatological and
XX neuroprotective activities. (I) can be used as a therapeutic or
XX prophylactic agent as well as for screening purposes. (I) is useful for
XX diagnosing diseases characterised by dysfunction of their associated
XX protein of interest, for identifying normal or abnormal proteins of
XX interest, as a part of diagnostic kit to detect the presence of their
XX proteins of interest in a biological sample. Additionally, (I) is useful
XX for treating inflammatory and autoimmune diseases, tumour growth, cancer,
XX rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,
XX infertility, and neurological degenerative diseases. (I), comprising EPO-
XX mimetic compounds are useful for treating disorders characterised by low
XX red blood cell levels such as anaemia. The TPO-mimetic comprising
XX compounds are useful for treating conditions that involve an existing
XX megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet
XX deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic
XX tumour which result in thrombocytopaenia, systemic lupus erythematosus,
```

CC and Panconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777
 CC represent amino acid and nucleic acid sequences used in the
 CC exemplification of the present invention

XX
 SQ Sequence 12 AA;

Query Match 98.5%; Score 66; DB 5; Length 12;
 Best Local Similarity 90.0%; Pred. No. 0.71;
 Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPYPPPPPIP 10
 Db 1 PPPYPPPPVP 10
 |||||:|

RESULT 15
 ADJ73404
 ID ADJ73404 standard; peptide; 12 AA.
 XX
 AC ADJ73404;
 XX
 DT 06-MAY-2004 (first entry)
 XX
 DE SH3 antagonist peptide sequence SeqID 859.
 XX
 KW mimetic; CDR mimetibody; gene therapy; transgenic; immune;
 KW cardiovascular; infectious; malignant; neurologic disease; anaemia;
 KW immunomodulator; cardiant; antimicrobial; cytostatic; neuroprotective;
 KW SH3.
 XX
 OS Synthetic.
 XX
 PN WO2003084477-A2.
 XX
 PD 16-OCT-2003.
 XX
 PF 24-MAR-2003; 2003WO-US009139.
 XX
 PR 29-MAR-2002; 2002US-0368791P.
 XX
 PA (CENZ) CENTOCOR INC.
 XX
 PI Heavner GA, Knight DM, Scallion BJ, Ghrayeb J;
 XX WPI; 2003-804237/75.
 XX
 DR New CDR mimetibody comprising a portion of a heavy or light chain
 XX variable region comprising human framework or ligand binding region,
 PT useful for preparing a composition for treating e.g., immune,
 PT cardiovascular or neurologic disease.
 XX
 PS Disclosure; SEQ ID NO 859; 97pp; English.
 XX
 CC This invention relates to novel mammalian CDR mimetibodies, specific
 CC portions or variants thereof. Specifically, it refers to an antibody
 CC fragment where a protein has been inserted into, or replaces a portion
 CC of, one or more CDR regions, such that each CDR mimetibody comprises at
 CC least one portion of a heavy chain or light chain variable region, which
 CC itself comprises at least one human framework region and at least one
 CC ligand binding region (LBR). The present invention describes human
 CC mimetibodies, including modified immunoglobulins and cleavage products
 CC that can be useful in gene therapy and the generation of transgenic
 CC plants and animals. Furthermore, the CDR mimetibody is useful for
 CC preparing compositions for modulating, treating or reducing the symptoms
 CC of immune, cardiovascular, infectious, malignant and/or neurologic
 CC diseases, as well as anaemia. Accordingly, they exhibit immunomodulator,
 CC cardiant, antimicrobial, cytostatic and neuroprotective activities. This
 CC peptide sequence is an SH3 antagonist peptide sequence used to make a
 CC mimetibody of the invention.

XX
 SQ Sequence 12 AA;

Query Match 98.5%; Score 66; DB 7; Length 12;

Best Local Similarity 90.0%; Pred. No. 0.71;
 Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPYPPPPPIP 10
 Db 1 PPPYPPPPVP 10
 |||||:|

RESULT 16
 ADJ53038
 ID ADJ53038 standard; peptide; 12 AA.
 XX
 AC ADJ53038;
 XX
 DT 06-MAY-2004 (first entry)
 XX
 DE CH1 deleted mimetibody-related peptide SeqID859.
 XX
 KW CH1 deleted mimetibody; immunosuppressive; cardiovascular; cardiant;
 KW hypotensive; neuroprotective; nontropic; antibacterial; virucide;
 KW fungicide; gene therapy; immune disorder; cardiovascular disease;
 KW arrhythmia; hypertension; heart failure; neurodegenerative;
 KW multiple sclerosis; dementia; Alzheimer's disease; anaemia;
 KW cancerous condition; infectious disease; bacterial infection;
 KW viral infection; fungal infection.
 XX
 OS Unidentified.
 OS Synthetic.
 XX
 PH Key Location/Qualifiers
 FT Misc-difference 1..12
 FT /note= "All Xaa's in this sequence are unidentified amino
 FT acids"
 XX
 PN WO2004002417-A2.
 XX
 PD 08-JAN-2004.
 XX
 PF 27-JUN-2003; 2003WO-US020347.
 XX
 PR 28-JUN-2002; 2002US-0392431P.
 XX
 PA (CENZ) CENTOCOR INC.
 XX
 PI Heavner GA, Knight DM, Ghrayeb J, Scallion BJ, Nesspor TC;
 PI Kutloski KA;
 XX WPI; 2004-082870/08.
 XX
 DR New CH1-deleted mimetibody polypeptides and nucleic acids, useful for
 XX modulating, treating, alleviating, preventing an immune, cardiovascular,
 PT or neurodegenerative disease or disorder, anemia, cancer, or infectious
 PT diseases.
 XX
 PS Claim 3; SEQ ID NO 859; 129pp; English.
 XX
 CC This invention relates to CH1 deleted mimetibodies (and the DNA sequences
 CC which encode them), compositions, methods and uses. The invention may be
 CC useful for the development of compounds with an immunosuppressive,
 CC cardiovascular, cardiant, hypotensive, neuroprotective, nontropic,
 CC antibacterial, virucide or fungicide activity. In addition, the disclosed
 CC sequences may prove useful for gene therapy. The CH1-deleted mimetibody
 CC is useful for diagnosing or treating a disease condition in a cell,
 CC tissue, organ or animal, specifically for modulating, treating,
 CC alleviating, preventing the incidence or reducing the symptoms of an
 CC immune, cardiovascular (for example arrhythmia, hypertension or heart
 CC failure), or neurodegenerative (for example multiple sclerosis, dementia
 CC or Alzheimer's disease) diseases or disorders, anaemia, cancerous
 CC conditions, or infectious diseases (for example bacterial, viral or
 CC fungal infection). The present sequence is that of a peptide which may be
 CC used during the creation of a mimetibody of the invention.

XX
 SQ Sequence 12 AA;

Query Match 98.5%; Score 66; DB 8; Length 12;
 Best Local Similarity 90.0%; Pred. No. 0.71;
 Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPYPPPP 10
 |||||
 DB 1 PPPYPPPP 10
 |||||

RESULT 17
 ADJ51999
 ID ADJ51999 standard; peptide; 12 AA.
 AC ADJ51999;
 DT 06-MAY-2004 (first entry)
 DE CH1 deleted mimetibody-related peptide SeqID859.
 XX CH1 deleted mimetibody; osteopathic; cardiovascular-Gen;
 KW dermatological-Gen; auditory; endocrine-Gen; gastrointestinal-Gen;
 KW gynaecological-Gen; hepatotropic; haemostatic; immunomodulator;
 KW anti-allergic; muscular-Gen; cytostatic; anti-inflammatory; neuroleptic;
 KW ophthalmological; nephrotropic; respiratory-Gen; tumour necrosis factor;
 KW TNF; cytokine; bone disorder; joint disorder; cardiovascular disorder;
 KW dental disorder; oral disorder; dermatological disorder; ear disorder;
 KW nose disorder; throat disorder; endocrine disorder; metabolic disorder;
 KW gastrointestinal disorder; gynaecological disorder; hepatic disorder;
 KW obstetric disorder; haematologic disorder; immunological disorder;
 KW allergic disorder; infectious disorder; musculoskeletal disorder;
 KW oncological disorder; neurological disorder; nutritional disorder;
 KW ophthalmologic disorder; pediatric disorder; psychiatric disorder;
 KW renal disorder; pulmonary disorder.
 XX Unidentified.
 OS Synthetic.
 PH Key Location/Qualifiers
 FT Misc-difference 1..12
 FT /note= "All Xaa's in this sequence are unidentified amino
 FT acids"
 XX WO2004002424-A2.
 XX 08-JAN-2004.
 XX 30-JUN-2003; 2003WO-US020495.
 XX 28-JUN-2002; 2002US-0392431P.
 PR 19-SEP-2002; 2002US-0412144P.
 XX (CENZ) CENTOCOR INC.
 PA Heavner GA, Knight DM, Ghayeb J, Scallion BJ, Nesspor TC;
 PI Kutloski KA;
 XX WPI; 2004-082872/08.
 XX New CH1 deleted mimetibody polypeptide and nucleic acid, useful for
 PT diagnosing, preventing or treating cardiovascular, dermatologic,
 PT endocrine, gastrointestinal, gynecologic, infectious, neurologic and
 PT nutritional disorders.
 XX Claim 15; SEQ ID NO 859; 123pp; English.
 XX This invention relates to CH1 deleted mimetibodies (and the DNA sequences
 CC which encode them), compositions, methods and uses. The invention may be
 CC useful for the development of compounds with an osteopathic, CC
 CC cardiovascular-Gen, dermatological-Gen, auditory, endocrine-Gen,
 CC gastrointestinal-Gen, gynaecological-Gen, hepatotropic, haemostatic,
 CC immunomodulator, anti-allergic, muscular-Gen, cytostatic,
 CC anti-inflammatory, neuroleptic, ophthalmological, nephrotropic or

CC respiratory-Gen activity acting as a tumour necrosis factor (TNF)-
 CC modulator or cytokine-agonist. The methods and compositions of the
 CC present invention are useful for the diagnosis, prevention and/or
 CC treatment of diseases or conditions associated with aberrant expression
 CC or activity of the CH1 deleted mimetibody, such as a bone or joint,
 CC cardiovascular, dental or oral, dermatological, ear, nose or throat,
 CC endocrine, metabolic, gastrointestinal, gynaecological, hepatic,
 CC obstetric, haematologic, immunological, allergic, infectious,
 CC musculoskeletal, oncological, neurological, nutritional, ophthalmologic,
 CC pediatric, psychiatric, renal or pulmonary disorders. The present
 CC sequence is that of a peptide which may be used during the creation of a
 CC mimetibody of the invention.
 XX
 SQ Sequence 12 AA;
 Query Match 98.5%; Score 66; DB 8; Length 12;
 Best Local Similarity 90.0%; Pred. No. 0.71;
 Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPYPPPP 10
 |||||
 DB 1 PPPYPPPP 10
 |||||

RESULT 18
 AAR93315
 ID AAR93315 standard; peptide; 10 AA.
 AC AAR93315;
 XX 24-APR-1996 (first entry)
 DT ABL protein tyrosine kinase derived peptide #2.
 DE SH3 ligand; SH3 binding agent; biased phage library;
 KW recognition sequence; src SH3 domain; Paget's disease; restenosis;
 KW rheumatoid arthritis; gout; neutrophil oxidase; tyrosine kinase; p47;
 KW p67; complex; chronic myelogenous leukaemia; cancer.
 XX Synthetic.
 OS WO9524419-A1.
 XX 14-SEP-1995.
 XX 13-MAR-1995; 95WO-US003208.
 XX 11-MAR-1994; 94US-00209835.
 PR 06-JAN-1995; 95US-00369832.
 XX (ARIA-) ARIAD PHARM INC.
 PA Rickles RJ, Brugge JS, Botfield MC, Zoller MJ;
 PI WPI; 1995-328231/42.
 XX Identification of peptide(s) binding specifically to SH3 domains - for
 PT use in inhibiting interactions mediated by SH3 domains in treatment of
 PT e.g. osteoporosis and cancer.
 XX Disclosure; Fig 1; 74pp; English.
 XX The sequences given in AAR93272-342 represent peptides which are SH3
 CC ligands/SH3 binding agents. They represent a biased phage library which
 CC comprises six random amino acids flanking the tetrapeptide -PPYP which
 CC was identified as a recognition sequence for the src SH3 domain. These
 CC sequences were identified using the method of the invention. The method
 CC comprises contacting the SH3 domain with a mixture of peptides under
 CC conditions permitting a ligand to bind to an SH3 domain to form a
 CC complex. Any unbound peptides are removed and the complexed peptide
 CC ligands are dissociated from the complexes. The selected peptides are
 CC enriched by re-contacting them with the SH3 domain and then candidates
 CC which bind to the SH3 domain are detected. The isolated SH3 binding

CC peptides may be used in the diagnosis, prevention and treatment of
 CC conditions or diseases resulting from cellular processes mediated by an
 CC SH3-based interaction. Such diseases include Paget's disease. Other
 CC conditions treatable with these peptides include restenosis, rheumatoid
 CC arthritis, gout and other problems in which an SH3 of neutrophil oxidase
 CC p47 and p67 complex is implicated, etc
 XX

SQ Sequence 10 AA;

Query Match 89.6%; Score 60; DB 2; Length 10;

Best Local Similarity 100.0%; Pred. No. 2.6;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 PPYP PPPPIP 10

Db 2 PPYP PPPPIP 10

RESULT 19

AAR93318

ID AAR93318 standard; peptide; 10 AA.

XX AC AAR93318;

XX DT 24-APR-1996 (first entry)

XX DE ABL protein tyrosine kinase derived peptide #5.

XX KW SH3 ligand; SH3 binding agent; biased phage library;

XX KW recognition sequence; src SH3 domain; Paget's disease; restenosis;

XX KW rheumatoid arthritis; gout; neutrophil oxidase; tyrosine kinase; p47;

XX KW p67; complex; chronic myelogenous leukaemia; cancer.

XX OS Synthetic.

XX PN WO9524419-A1.

XX PD 14-SEP-1995.

XX PF 13-MAR-1995; 95WO-US003208.

XX PR 11-MAR-1994; 94US-00209835.

XX PR 06-JAN-1995; 95US-00369832.

XX PA (ARIA-) ARIAD PHARM INC.

XX PI Rickles RJ, Brugge JS, Botfield MC, Zoller MJ;

XX DR WPI; 1995-328231/42.

XX PT Identification of peptide(s) binding specifically to SH3 domains - for
 XX use in inhibiting interactions mediated by SH3 domains in treatment of
 XX e.g. osteoporosis and cancer.

XX PS Disclosure; Fig 1; 74pp; English.

XX CC The sequences given in AAR93272-342 represent peptides which are SH3
 CC ligands/SH3 binding agents. They represent a biased phage library which
 CC comprises six random amino acids flanking the tetrapeptide -PPIP which
 CC was identified as a recognition sequence for the src SH3 domain. These
 CC sequences were identified using the method of the invention. The method
 CC comprises contacting the SH3 domain with a mixture of peptides under
 CC conditions permitting a ligand to bind to an SH3 domain to form a
 CC complex. Any unbound peptides are removed and the complexed peptide
 CC ligands are dissociated from the complexes. The selected peptides are
 CC enriched by re-contacting them with the SH3 domain and then candidates
 CC which bind to the SH3 domain are detected. The isolated SH3 binding
 CC peptides may be used in the diagnosis, prevention and treatment of
 CC conditions or diseases resulting from cellular processes mediated by an
 CC SH3-based interaction. Such diseases include Paget's disease. Other
 CC conditions treatable with these peptides include restenosis, rheumatoid
 CC arthritis, gout and other problems in which an SH3 of neutrophil oxidase
 CC p47 and p67 complex is implicated, etc

XX SQ Sequence 10 AA;

Query Match 88.1%; Score 59; DB 2; Length 10;

Best Local Similarity 90.0%; Pred. No. 3.4;

Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPYP PPPPIP 10

Db 1 PPYP PPPPIP 10

RESULT 20

AAR93319

ID AAR93319 standard; peptide; 10 AA.

XX AC AAR93319;

XX DT 24-APR-1996 (first entry)

XX DE ABL protein tyrosine kinase derived peptide #6.

XX KW SH3 ligand; SH3 binding agent; biased phage library;

XX KW recognition sequence; src SH3 domain; Paget's disease; restenosis;

XX KW rheumatoid arthritis; gout; neutrophil oxidase; tyrosine kinase; p47;

XX KW p67; complex; chronic myelogenous leukaemia; cancer.

XX OS Synthetic.

XX PN WO9524419-A1.

XX PD 14-SEP-1995.

XX PF 13-MAR-1995; 95WO-US003208.

XX PR 11-MAR-1994; 94US-00209835.

XX PR 06-JAN-1995; 95US-00369832.

XX PA (ARIA-) ARIAD PHARM INC.

XX PI Rickles RJ, Brugge JS, Botfield MC, Zoller MJ;

XX DR WPI; 1995-328231/42.

XX PT Identification of peptide(s) binding specifically to SH3 domains - for
 XX use in inhibiting interactions mediated by SH3 domains in treatment of
 XX e.g. osteoporosis and cancer.

XX PS Disclosure; Fig 1; 74pp; English.

XX CC The sequences given in AAR93272-342 represent peptides which are SH3
 CC ligands/SH3 binding agents. They represent a biased phage library which
 CC comprises six random amino acids flanking the tetrapeptide -PPIP which
 CC was identified as a recognition sequence for the src SH3 domain. These
 CC sequences were identified using the method of the invention. The method
 CC comprises contacting the SH3 domain with a mixture of peptides under
 CC conditions permitting a ligand to bind to an SH3 domain to form a
 CC complex. Any unbound peptides are removed and the complexed peptide
 CC ligands are dissociated from the complexes. The selected peptides are
 CC enriched by re-contacting them with the SH3 domain and then candidates
 CC which bind to the SH3 domain are detected. The isolated SH3 binding
 CC peptides may be used in the diagnosis, prevention and treatment of
 CC conditions or diseases resulting from cellular processes mediated by an
 CC SH3-based interaction. Such diseases include Paget's disease. Other
 CC conditions treatable with these peptides include restenosis, rheumatoid
 CC arthritis, gout and other problems in which an SH3 of neutrophil oxidase
 CC p47 and p67 complex is implicated, etc

XX SQ Sequence 10 AA;

Query Match 88.1%; Score 59; DB 2; Length 10;

Best Local Similarity 90.0%; Pred. No. 3.4;

Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY      1 PYPYPPPPIP 10
DB      1 PPAYPPPPIP 10

RESULT 21
AAR93374
ID AAR93374 standard; peptide; 11 AA.
XX
AC AAR93374;
XX
DT 24-APR-1996 (first entry)
XX
DE ABL protein tyrosine kinase derived peptide #6.
XX
KW SH3 ligand; SH3 binding agent; biased phage library;
KW recognition sequence; src SH3 domain; Paget's disease; restenosis;
KW rheumatoid arthritis; gout; neutrophil oxidase; tyrosine kinase; p47;
KW p67; complex; chronic myelogenous leukaemia; cancer.
XX
OS Synthetic.
XX
PN WO9524419-A1.
XX
PD 14-SEP-1995.
XX
PF 13-MAR-1995; 95WO-US003208.
XX
PR 11-MAR-1994; 94US-00209835.
XX
PR 06-JAN-1995; 95US-00369832.
XX
PA (ARIA-) ARIAD PHARM INC.
XX
PI Rickles RJ, Brugge JS, Botfield MC, Zoller MJ;
XX WPI; 1995-328231/42.
XX
PT Identification of peptide(s) binding specifically to SH3 domains - for
PT use in inhibiting interactions mediated by SH3 domains in treatment of
PT e.g. osteoporosis and cancer.
XX
PS Disclosure; Fig 2; 74pp; English.
XX
CC The sequences given in AAR93369-76 represent peptides which are SH3
CC ligands/SH3 binding agents. They represent a biased phage library which
CC comprises six random amino acids flanking the hexapeptide PYPYPP- which
CC was identified as a recognition sequence for the src SH3 domain. These
CC sequences were identified using the method of the invention. The method
CC comprises contacting the SH3 domain with a mixture of peptides under
CC conditions permitting a ligand to bind to an SH3 domain to form a
CC complex. Any unbound peptides are removed and the complexed peptide
CC ligands are dissociated from the complexes. The selected peptides are
CC enriched by re-contacting them with the SH3 domain and then candidates
CC which bind to the SH3 domain are detected. The isolated SH3 binding
CC peptides may be used in the diagnosis, prevention and treatment of
CC conditions or diseases resulting from cellular processes mediated by an
CC SH3-based interaction. Such diseases include Paget's disease. Other
CC conditions treatable with these peptides include restenosis, rheumatoid
CC arthritis, gout and other problems in which an SH3 of neutrophil oxidase
CC p47 and p67 complex is implicated, etc
XX
SQ Sequence 11 AA;

Query Match      88.1%; Score 59; DB 2; Length 11;
Best Local Similarity 88.9%; Pred. NO. 3.6;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      2 PYPYPPPPIP 10
DB      1 PPYPPPPVP 9

RESULT 22
AAR93376
ID AAR93376 standard; peptide; 12 AA.
XX
AC AAR93376;
XX
DT 24-APR-1996 (first entry)
XX
DE ABL protein tyrosine kinase derived peptide #8.
XX
KW SH3 ligand; SH3 binding agent; biased phage library;
KW recognition sequence; src SH3 domain; Paget's disease; restenosis;
KW rheumatoid arthritis; gout; neutrophil oxidase; tyrosine kinase; p47;
KW p67; complex; chronic myelogenous leukaemia; cancer.
XX
OS Synthetic.
XX
PN WO9524419-A1.
XX
PD 14-SEP-1995.
XX
PF 13-MAR-1995; 95WO-US003208.
XX
PR 11-MAR-1994; 94US-00209835.
XX
PR 06-JAN-1995; 95US-00369832.
XX
PA (ARIA-) ARIAD PHARM INC.
XX
PI Rickles RJ, Brugge JS, Botfield MC, Zoller MJ;
XX WPI; 1995-328231/42.
XX
PT Identification of peptide(s) binding specifically to SH3 domains - for
PT use in inhibiting interactions mediated by SH3 domains in treatment of
PT e.g. osteoporosis and cancer.
XX
PS Disclosure; Fig 2; 74pp; English.
XX
CC The sequences given in AAR93369-76 represent peptides which are SH3
CC ligands/SH3 binding agents. They represent a biased phage library which
CC comprises six random amino acids flanking the hexapeptide PYPYPP- which
CC was identified as a recognition sequence for the src SH3 domain. These
CC sequences were identified using the method of the invention. The method
CC comprises contacting the SH3 domain with a mixture of peptides under
CC conditions permitting a ligand to bind to an SH3 domain to form a
CC complex. Any unbound peptides are removed and the complexed peptide
CC ligands are dissociated from the complexes. The selected peptides are
CC enriched by re-contacting them with the SH3 domain and then candidates
CC which bind to the SH3 domain are detected. The isolated SH3 binding
CC peptides may be used in the diagnosis, prevention and treatment of
CC conditions or diseases resulting from cellular processes mediated by an
CC SH3-based interaction. Such diseases include Paget's disease. Other
CC conditions treatable with these peptides include restenosis, rheumatoid
CC arthritis, gout and other problems in which an SH3 of neutrophil oxidase
CC p47 and p67 complex is implicated, etc
XX
SQ Sequence 12 AA;

Query Match      88.1%; Score 59; DB 2; Length 12;
Best Local Similarity 90.0%; Pred. NO. 3.9;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 PYPYPPPPIP 10
DB      1 PPYPYPPDIP 10

RESULT 23
AAR93317
ID AAR93317 standard; peptide; 10 AA.
XX
AC AAR93317;
XX

```


DT 24-APR-1996 (first entry)
 XX ABL protein tyrosine kinase derived peptide #4.
 DE
 XX
 KW SH3 ligand; SH3 binding agent; biased phage library;
 KW recognition sequence; src SH3 domain; Paget's disease; restenosis;
 KW rheumatoid arthritis; gout; neutrophil oxidase; tyrosine kinase; p47;
 KW p67; complex; chronic myelogenous leukaemia; cancer.
 XX
 OS Synthetic.
 XX
 PN WO9524419-A1.
 XX
 PD 14-SEP-1995.
 XX
 PF 13-MAR-1995; 95WO-US003208.
 XX
 PR 11-MAR-1994; 94US-00209835.
 PR 06-JAN-1995; 95US-00369832.
 XX
 PA (ARIA-) ARIAD PHARM INC.
 XX
 PI Rickles RJ, Brugge JS, Botfield MC, Zoller MJ;
 XX WPI; 1995-328231/42.
 DR
 XX
 PT Identification of peptide(s) binding specifically to SH3 domains - for
 PT use in inhibiting interactions mediated by SH3 domains in treatment of
 PT e.g. osteoporosis and cancer.
 XX
 PS Disclosure; Fig 1; 74pp; English.
 XX
 CC The sequences given in AAR93272-342 represent peptides which are SH3
 CC ligands/SH3 binding agents. They represent a biased phage library which
 CC comprises six random amino acids flanking the tetrapeptide -ppip which
 CC was identified as a recognition sequence for the src SH3 domain. These
 CC sequences were identified using the method of the invention. The method
 CC comprises contacting the SH3 domain with a mixture of peptides under
 CC conditions permitting a ligand to bind to an SH3 domain to form a
 CC complex. Any unbound peptides are removed and the complexed peptide
 CC ligands are dissociated from the complexes. The selected peptides are
 CC enriched by re-contacting them with the SH3 domain and then candidates
 CC which bind to the SH3 domain are detected. The isolated SH3 binding
 CC peptides may be used in the diagnosis, prevention and treatment of
 CC conditions or diseases resulting from cellular processes mediated by an
 CC SH3-based interaction. Such diseases include Paget's disease. Other
 CC conditions treatable with these peptides include restenosis, rheumatoid
 CC arthritis, gout and other problems in which an SH3 of neutrophil oxidase
 CC p47 and p67 complex is implicated, etc
 XX
 SQ Sequence 10 AA;
 Query Match 86.6%; Score 58; DB 2; Length 10;
 Best Local Similarity 90.0%; Pred. No. 4.3;
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 P P P Y P P P P I P 10
 Db || || || || || || || || || ||
 1 P P H Y P P P P I P 10
 RESULT 25
 AAB98984
 ID AAB98984 standard; peptide; 10 AA.
 XX
 AC AAB98984;
 XX
 DT 20-AUG-2001 (first entry)
 XX
 DE Yeast ORF YDL117W SH3 domain peptide ligand.
 XX
 KW Yeast; ORF YDL117W; protein 3-dimensional modelling; 3-d modelling;
 KW protein structure analysis; ligand.
 XX
 OS Unidentified.
 XX
 PN WO200135255-A2.
 XX
 PD 17-MAY-2001.

DT 24-APR-1996 (first entry)
 XX ABL protein tyrosine kinase derived peptide #4.
 DE
 XX
 KW SH3 ligand; SH3 binding agent; biased phage library;
 KW recognition sequence; src SH3 domain; Paget's disease; restenosis;
 KW rheumatoid arthritis; gout; neutrophil oxidase; tyrosine kinase; p47;
 KW p67; complex; chronic myelogenous leukaemia; cancer.
 XX
 OS Synthetic.
 XX
 PN WO9524419-A1.
 XX
 PD 14-SEP-1995.
 XX
 PF 13-MAR-1995; 95WO-US003208.
 XX
 PR 11-MAR-1994; 94US-00209835.
 PR 06-JAN-1995; 95US-00369832.
 XX
 PA (ARIA-) ARIAD PHARM INC.
 XX
 PI Rickles RJ, Brugge JS, Botfield MC, Zoller MJ;
 XX WPI; 1995-328231/42.
 DR
 XX
 PT Identification of peptide(s) binding specifically to SH3 domains - for
 PT use in inhibiting interactions mediated by SH3 domains in treatment of
 PT e.g. osteoporosis and cancer.
 XX
 PS Disclosure; Fig 1; 74pp; English.
 XX
 CC The sequences given in AAR93272-342 represent peptides which are SH3
 CC ligands/SH3 binding agents. They represent a biased phage library which
 CC comprises six random amino acids flanking the tetrapeptide -ppip which
 CC was identified as a recognition sequence for the src SH3 domain. These
 CC sequences were identified using the method of the invention. The method
 CC comprises contacting the SH3 domain with a mixture of peptides under
 CC conditions permitting a ligand to bind to an SH3 domain to form a
 CC complex. Any unbound peptides are removed and the complexed peptide
 CC ligands are dissociated from the complexes. The selected peptides are
 CC enriched by re-contacting them with the SH3 domain and then candidates
 CC which bind to the SH3 domain are detected. The isolated SH3 binding
 CC peptides may be used in the diagnosis, prevention and treatment of
 CC conditions or diseases resulting from cellular processes mediated by an
 CC SH3-based interaction. Such diseases include Paget's disease. Other
 CC conditions treatable with these peptides include restenosis, rheumatoid
 CC arthritis, gout and other problems in which an SH3 of neutrophil oxidase
 CC p47 and p67 complex is implicated, etc
 XX
 SQ Sequence 10 AA;
 Query Match 86.6%; Score 58; DB 2; Length 10;
 Best Local Similarity 90.0%; Pred. No. 4.3;
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 P P P Y P P P P I P 10
 Db || || || || || || || || || ||
 1 P P P H P P P P I P 10
 RESULT 24
 AAR93323
 ID AAR93323 standard; peptide; 10 AA.
 XX
 AC AAR93323;
 XX
 DT 24-APR-1996 (first entry)
 XX
 DE ABL protein tyrosine kinase derived peptide #10.
 XX
 KW SH3 ligand; SH3 binding agent; biased phage library;
 KW recognition sequence; src SH3 domain; Paget's disease; restenosis;
 KW

XX 08-NOV-2000; 2000MO-US030753.
 XX
 XX 09-NOV-1999; 99US-00437738.
 XX
 XX (UYRQ) UNIV ROCKEFELLER.
 XX
 XX Sali A, Sanchez R, Melo F;
 XX WPI; 2001-335968/35.
 XX
 XX Three dimensional model of protein is generated by searching database(s)
 PT to find proteins similar to selected protein amino acid sequence and
 PT selecting best model based on selected proteins.
 XX
 XX Example 1; Page 51; 124pp; English.
 XX
 XX The present invention describes a computerised process of generating a
 CC three dimensional model of a protein, involving searching databases to
 CC find proteins having more than a preset similarity to the protein of
 CC interest. The process can be used to model protein structures. The
 CC present sequence is a peptide fragment used to demonstrate the invention
 XX
 XX Sequence 10 AA;
 SQ
 Query Match 86.6%; Score 58; DB 4; Length 10;
 Best Local Similarity 80.0%; Pred. No. 4.3;
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 PPPPPPPPIP 10
 Db || |||||:
 1 PPAYPPPPVP 10

RESULT 26
 ADK70159
 ID ADK70159 standard; peptide; 10 AA.
 XX
 XX ADK70159;
 XX
 XX 06-MAY-2004 (first entry)
 XX
 XX Human SH3 binding protein 2 SH3 ligand motif.
 DE
 XX
 XX exon; mutant; osteopathic; cytostatic; gene therapy; SH3-binding protein;
 KW SH3BP2; mutation; diagnosis; bone homeostasis; cherubism; bone tumor.
 XX
 XX Homo sapiens.
 OS
 XX WO2003025197-A2.
 PN
 XX 27-MAR-2003.
 PD
 XX 01-FEB-2002; 2002MO-US019164.
 PF
 XX 02-FEB-2001; 2001US-0266129P.
 PR
 XX (HARD) HARVARD COLLEGE.
 PA (FORS-) FORSYTH DENTAL INFIRMARY FOR CHILDREN.
 XX
 XX Tiziani V, Reichenberger E, Ueki Y, Olsen BR;
 PI WPI; 2003-371820/35.
 XX
 XX New mutant SH3BP2 nucleic acid molecule or polypeptide, useful for
 PT diagnosing and treating disorders of bone homeostasis, such as cherubism
 PT or bone tumor.
 XX
 XX Example 1; SEQ ID NO 103; 70pp; English.
 PS
 XX The invention relates to an isolated nucleic acid molecule comprising a
 CC mutant SH3-binding protein and its encoding DNA, nucleic acid molecules
 CC which hybridize under stringent conditions to it, a nucleic acid

CC molecules that differ from it in codon sequence due to the degeneracy of
 CC the genetic code; and complements of these, provided that the nucleic
 CC acid molecule is not a human 'Wt' SH3BP2 full length sequence. The mutant
 CC is especially selected from a genetic mutant domain; 'H' family mutation;
 CC 'K' family mutation; 'A,B' family mutation; 'C, P, J, M, O' family
 CC mutation; 'L' family mutation; 'G' family mutation; 'N' family mutation
 CC or genomic mutant SH3BP2), any of the mutant SH3BP2 exon 9 sequences and
 CC a genomic mutant SH3BP2 nucleic acid molecule. The nucleic acid molecule
 CC and polypeptide are useful for diagnosing and treating disorders of bone
 CC homeostasis, such as cherubism or a bone tumor. This sequence corresponds
 CC to the wild type SH3-binding protein ligand motif.
 XX
 XX Sequence 10 AA;
 SQ
 Query Match 86.6%; Score 58; DB 7; Length 10;
 Best Local Similarity 80.0%; Pred. No. 4.3;
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 PPPPPPPPIP 10
 Db || |||||:
 1 PPAYPPPPVP 10

RESULT 27
 ABG75623
 ID ABG75623 standard; peptide; 11 AA.
 XX
 XX ABG75623;
 AC
 XX 25-APR-2003 (first entry)
 DT
 XX Human Abelson protein tyrosine kinase ligand #2.
 DE
 XX
 XX Abelson protein tyrosine kinase; Abl; SH2; SH3; ligand; protein ligation;
 KW intein-chitin binding domain; CBD; N-terminal cysteine;
 KW unoxidised sulphhydryl side chain; protein-chip; semi-synthetic protein.
 XX
 XX Synthetic.
 OS
 XX
 XX Key Location/Qualifiers
 PH Modified-site 11
 FT /label= OTHER
 FT /note= "C-terminal amide"
 XX
 XX US2002151006-A1.
 PN
 XX 17-OCT-2002.
 PD
 XX 12-JUL-2001; 2001US-00904117.
 PF
 XX 13-NOV-1997; 97US-0065391P.
 PR 24-JUL-1998; 98US-0093990P.
 PR 13-NOV-1998; 98US-00191890.
 XX
 XX (MUIR/) MUIR T W.
 PA (COLE/) COLE P A.
 PA (FRIE/) FRIEDMAN J M.
 PA (SOND/) SONDHI D.
 PA (SEVE/) SEVERINOV K.
 XX
 XX Muir TW, Cole PA, Friedman JM, Sondhi D, Severinov K;
 PI WPI; 2003-238034/23.
 XX
 XX Clearing recombinantly expressed proteins bound to intein-chitin binding
 PT domains, and ligating to peptides with N-terminal cysteines having
 PT unoxidized sulphhydryl side chain, useful for producing semi-synthetic
 PT proteins.
 XX
 XX Example 3; Page 15; 35pp; English.
 PS
 XX The invention relates to cleaving a recombinantly expressed protein bound
 CC to an intein-chitin binding domain (CBD) and ligating the cleaved protein
 CC

CC to a peptide with an N-terminal cysteine having an unoxidised sulphhydryl
 CC side chain, by contacting the bound protein with a peptide in the
 CC presence of conjugated thiol to effect cleavage of the protein from the
 CC intein-CBD and production of a C-terminal thioester of the protein which
 CC undergoes intramolecular rearrangement to form amide bond linking the
 CC protein to the peptide. The method is used for linking recombinantly
 CC expressed protein bound to an intein-CBD to a peptide with N-terminal
 CC cysteine having an unoxidised sulphhydryl side chain. The method is
 CC useful for producing a protein-chip which is useful for diagnostic
 CC screening for a specific protein, antibody or antigen and for preparing
 CC semi-synthetic proteins of any size. The protein-chip is also useful for
 CC identifying presence of protein in a sample. The method is also useful
 CC for generating recombinant proteins e.g. cytotoxic recombinant proteins,
 CC recombinant proteins partially labeled with a detectable marker, in
 CC nuclear magnetic resonance (NMR) spectroscopy, where proteins used in the
 CC method are segmentally labeled, and for generating semi-synthetic
 CC proteins to facilitate two recombinant, folded proteins to be ligated
 CC together. Ligated SH2 and SH3 domains of the human Abl protein
 CC tyrosine kinase (Abl) were produced using the method of the invention.
 CC The present sequence is a ligand for Abl used to determine if the ligated
 CC domains have folded correctly
 XX
 SQ Sequence 11 AA;

Query Match 86.6%; Score 58; DB 6; Length 11;
 Best Local Similarity 80.0%; Pred. No. 4.7;
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 P P P Y P P P P P I P 10
 || || || || || || ||
 Db 1 P P A Y P P P P P V P 10

RESULT 28
 AAB12074
 ID AAB12074 standard; peptide; 40 AA.
 AC AAB12074;
 XX
 DT 17-JAN-2001 (first entry)
 XX
 DE SH3 binding domain from 3BP2.
 XX
 KW RB-interacting zinc finger; RIZ; retinoblastoma; heart disease;
 KW cell proliferation; cell differentiation; tissue repair;
 KW transcription regulator; breast cancer; gene therapy; melanoma;
 KW neuroblastoma; leukaemia; Parkinson's disease; Huntington's disease;
 KW Alzheimer's disease; paralysis; motor neurone disorder;
 KW SH3 binding domain.

XX Unidentified.
 XX
 XX US6069231-A.
 XX
 XX 30-MAY-2000.
 XX
 XX 18-AUG-1995; 95US-00516859.
 XX
 XX 18-AUG-1994; 94US-00292683.
 XX 06-MAR-1995; 95US-00399411.
 XX
 XX (LJOL-) LA JOLLA CANCER RES FOUND.
 XX
 XX Huang S;
 XX
 XX WPI; 2000-410879/35.

XX New PR domain peptides comprising amino acid sequences from, for example
 XX retinoblastoma-interacting zinc finger, or egl-43 proteins, for
 XX regulating gene transcription and controlling cell proliferation and
 XX differentiation.
 XX
 XX Disclosure; Fig 2; 91pp; English.

XX The present sequence is the SH3 binding domain from 3BP2 protein. A SH3
 CC binding domain is also found in rat RIZ (AAB12028). RIZ is retinoblastoma
 CC (Rb)-interacting zinc finger protein. RIZ is a nuclear phosphoprotein
 CC that acts as a cell differentiation factor. RIZ can modulate cell growth
 CC by binding to Rb protein, which is involved in regulating cell
 CC proliferation. In addition, RIZ can act to regulate transcription. RIZ
 CC functions to maintain cells in the G1 phase of the cell cycle, by
 CC interacting with Rb through the cr2 domain of RIZ. Rat RIZ protein
 CC contains a number of GTPase motifs (see AAB12037 to AAB12056 and AAB12099
 CC to AAB12104). RIZ protein is a PR domain protein and is present primarily
 CC in the cell nucleus. RIZ gene mutations may be implicated in various
 CC cancers such as melanoma, neuroblastoma, leukaemia and breast cancer, and
 CC so the RIZ gene may be used in gene therapy for these disorders. Since
 CC RIZ protein is implicated in cell cycle arrest, inhibition of RIZ
 CC activity may be useful in neurodegenerative disorder therapy e.g. for
 CC Parkinson's, Huntington's or Alzheimer's disease, paralysis or motor
 CC neurone disorders, or cardiac disorders e.g. heart disease, where the
 CC ability to induce neural/ cardiac tissue proliferation would be useful.
 CC The present sequence was used for sequence homology comparison
 XX
 SQ Sequence 40 AA;

Query Match 86.6%; Score 58; DB 3; Length 40;
 Best Local Similarity 80.0%; Pred. No. 13;
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 P P P Y P P P P P I P 10
 || || || || || || ||
 Db 1 P P A Y P P P P P V P 10

RESULT 29
 ADE84647
 ID ADE84647 standard; protein; 40 AA.
 AC ADE84647;
 XX
 DT 29-JAN-2004 (first entry)
 XX
 DE 3BP2 SH3 binding protein motif.
 XX
 KW cytosolic; tumor; gene therapy;
 KW retinoblastoma protein-interacting zinc protein; RIZ; neuroblastoma;
 KW melanoma; cardiac cell growth inducer; neuronal cell growth inducer;
 KW SH3 binding protein motif.
 XX
 OS Unidentified.
 XX
 XX US6469985-B1.
 XX
 XX 22-OCT-2002.
 XX
 XX 17-MAR-2000; 2000US-00528706.
 XX
 XX 18-AUG-1994; 94US-00292683.
 XX 06-MAR-1995; 95US-00399411.
 XX 18-AUG-1995; 95US-00516859.
 XX
 XX (BURN-) BURNHAM INST.
 XX
 XX Huang S;
 XX
 XX WPI; 2003-147106/14.
 XX
 XX Reducing the growth of tumor cell having mammalian retinoblastoma protein
 XX -interacting zinc finger proteins in a subject, comprises administering
 XX nucleic acid molecule encoding RIZ at or adjacent to the site of tumor.
 XX
 XX Disclosure; SEQ ID NO 60; 49pp; English.
 XX
 XX The invention describes a method of reducing growth of a tumour cell
 CC having a mutant mammalian retinoblastoma (Rb) protein-interacting zinc


```
RESULT 32
ADQ66376
ID ADQ66376 standard; protein; 536 AA.
XX AC ADQ66376;
XX AC ADQ66376;
XX DT 07-OCT-2004 (first entry)
XX DE
XX DE Novel human protein sequence #1349.
XX KW osteopathic; neuroprotective; nootropic; antiparkinsonian; cytostatic;
XX KW gene therapy; diagnostic marker; morbid state; osteoporosis;
XX KW neurological disease; Alzheimer's disease; Parkinson's disease; dementia;
XX KW cancer.
XX OS Homo sapiens.
XX PN EP1440981-A2.
XX PD 28-JUL-2004.
XX PF 21-JAN-2004; 2004EP-00001196.
XX PR 21-JAN-2003; 2003JP-00102206.
XX PR 09-MAY-2003; 2003JP-00131392.
XX PA (REAS-) RES ASSOC BIOTECHNOLOGY.
XX PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
XX PI Yamamoto J, Isono Y, Negai K, Irie R;
XX DR WPI; 2004-535376/52.
XX DR N-PSDB; ADQ64188.
XX PT Novel 2495 cDNA, useful for treating osteoporosis, neurological diseases,
XX PT Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.
XX PS Claim 1; SEQ ID NO 3537; 2449pp; English.
XX CC The invention relates to 2495 novel polynucleotides (I) and their encoded
XX CC polypeptides, sequences hybridizing to these nucleotides, sequences
XX CC encoding partial polypeptides and sequences having 70% or 90% identity to
XX CC the nucleotide and protein sequences. The nucleotides and polypeptides
XX CC are useful as diagnostic markers or therapeutic target for the diseases
XX CC or morbid states. They are also useful for treating osteoporosis,
XX CC neurological diseases, Alzheimer's diseases, Parkinson's diseases,
XX CC dementia and various cancers. This sequence corresponds to a protein
XX CC sequence of the invention.
XX SQ Sequence 536 AA;

Query Match 86.6%; Score 58; DB 8; Length 536;
Best Local Similarity 80.0%; Pred. No. 1.1e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPYPPPPPIP 10
Db 176 PPAYPPPPVP 185
||| |||||:|
||| |||||:|

RESULT 33
ADQ71301
ID ADC71301 standard; protein; 561 AA.
XX AC ADC71301;
XX DT 18-DEC-2003 (first entry)
XX DE
XX DE Human colon specific protein sequence DEX0235_129 (SeqID 129).
XX KW human; gene; ss; neoplastic colorectal; colon cancer;

Query Match 86.6%; Score 58; DB 8; Length 536;
Best Local Similarity 80.0%; Pred. No. 1.1e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPYPPPPPIP 10
Db 201 PPAYPPPPVP 210
||| |||||:|
||| |||||:|

RESULT 34
ADK70062
ID ADK70062 standard; protein; 561 AA.
XX AC ADK70062;
XX DT 06-MAY-2004 (first entry)
XX DE
XX DE Mutant human SH3 binding protein 2.
XX KW osteopathic; cytostatic; gene therapy; SH3-binding protein; SH3BP2;
XX KW mutation; diagnosis; bone homeostasis; cherubism; bone tumor; mutant;
XX KW mutein.
XX OS Homo sapiens.
XX OS Synthetic.
XX PN WO2003025197-A2.
XX PD 27-MAR-2003.
XX PF 01-FEB-2002; 2002WO-US019164.
XX PR 02-FEB-2001; 2001US-0266129P.

KW non-cancerous disease; gene therapy; transgenic; DEX0235_129.
XX Homo sapiens.
XX OS WO2003020934-A1.
XX PN 13-MAR-2003.
XX PD 29-AUG-2002; 2002WO-US027737.
XX PF 31-AUG-2001; 2001US-0316258P.
XX PR (DIAD-) DIADEXUS INC.
XX PA Sun Y, Liu C, Ghosh MG;
XX PI WPI; 2003-300891/29.
XX DR N-PSDB; ADC71368.
XX DR Novel colon specific polypeptides and nucleic acids, useful for
XX DR identifying, diagnosing, monitoring, staging, imaging and treating colon
XX DR cancer and non-cancerous disease states in colon tissue.
XX PS Claim 11; SEQ ID NO 129; 262pp; English.
XX CC This invention relates to novel nucleic acid molecules and the encoded
XX CC polypeptides, which are present in normal and neoplastic colorectal
XX CC cells. Specifically, it refers to antibodies of these colon specific
XX CC polypeptides, as well as antagonists and agonists thereof that can be
XX CC used to treat colon cancer and also non-cancerous diseases states of the
XX CC colon. The present invention describes methods useful for the diagnosis
XX CC and monitoring of colon cancer metastases in a patient, by determining
XX CC the concentration of these colon specific proteins in a patient sample.
XX CC Furthermore, they are also used for gene therapy purposes, the production
XX CC of transgenic animals and cells, as well as producing engineered colon
XX CC tissue for treatment and research. This polypeptide sequence is a human
XX CC colon specific protein sequence of the invention.
XX SQ Sequence 561 AA;

Query Match 86.6%; Score 58; DB 7; Length 561;
Best Local Similarity 80.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPYPPPPPIP 10
Db 201 PPAYPPPPVP 210
||| |||||:|
||| |||||:|
```

```
XX PA (HARD ) HARVARD COLLEGE.
XX PA (FORS-) FORSYTH DENTAL INFIRMARY FOR CHILDREN.
XX PI Tiziani V, Reichenberger E, Ueki Y, Olsen BR;
XX XX
XX DR WPI; 2003-371820/35.
XX DR N-PSDB; ADK70061.
XX XX
XX PT New mutant SH3BP2 nucleic acid molecule or polypeptide, useful for
XX PT diagnosing and treating disorders of bone homeostasis, such as cherubism
XX PT or bone tumor.
XX XX
XX PS Claim 9; SEQ ID NO 6; 70pp; English.
XX PS
XX CC The invention relates to an isolated nucleic acid molecule comprising a
XX CC mutant SH3-binding protein and its encoding DNA, nucleic acid molecules
XX CC which hybridize under stringent conditions to it, a nucleic acid
XX CC molecules that differ from it in codon sequence due to the degeneracy of
XX CC the genetic code; and complements of these, provided that the nucleic
XX CC acid molecule is not a human 'wt' SH3BP2 full length sequence. The mutant
XX CC is especially selected from a genetic mutant domain; 'H' family mutation;
XX CC 'K' family mutation; 'A,B' family mutation; 'C, F, J, M, O' family
XX CC mutation; 'L' family mutation; 'G' family mutation; 'N' family mutation
XX CC or genomic mutant SH3BP2), any of the mutant SH3BP2 exon 9 sequences and
XX CC a genomic mutant SH3BP2 nucleic acid molecule. The nucleic acid molecule
XX CC and polypeptide are useful for diagnosing and treating disorders of bone
XX CC homeostasis, such as cherubism or a bone tumor. This sequence corresponds
XX CC to the mutant full length SH3-binding protein.
XX SQ Sequence 561 AA;

  Query Match      86.6%; Score 58; DB 7; Length 561;
  Best Local Similarity 80.0%; Pred. No. 1.2e+02;
  Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPPYPPPPPIP 10
Db || |||||
  201 PPAYPPPPVP 210

RESULT 35
ADK70058
ID ADK70058 standard; protein; 561 AA.
AC ADK70058;
XX 06-MAY-2004 (first entry)
XX
XX DE Wild type human SH3 binding protein 2.
XX
XX KW osteopathic; cytostatic; gene therapy; SH3-binding protein; SH3BP2;
XX KW mutation; diagnosis; bone homeostasis; cherubism; bone tumor.
XX
XX OS Homo sapiens.
XX
XX PN WO2003025197-A2.
XX
XX PD 27-MAR-2003.
XX
XX PF 01-FEB-2002; 2002WO-US019164.
XX
XX PR 02-FEB-2001; 2001US-0266129P.
XX
XX PA (HARD ) HARVARD COLLEGE.
XX PA (FORS-) FORSYTH DENTAL INFIRMARY FOR CHILDREN.
XX PI Tiziani V, Reichenberger E, Ueki Y, Olsen BR;
XX XX
XX DR WPI; 2003-371820/35.
XX DR N-PSDB; ADK70057.
XX XX
XX PT New mutant SH3BP2 nucleic acid molecule or polypeptide, useful for
XX PT diagnosing and treating disorders of bone homeostasis, such as cherubism
XX PT or bone tumor.

  Query Match      86.6%; Score 58; DB 7; Length 561;
  Best Local Similarity 80.0%; Pred. No. 1.2e+02;
  Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPPYPPPPPIP 10
Db || |||||
  201 PPAYPPPPVP 210

RESULT 36
ADQ96506
ID ADQ96506 standard; protein; 561 AA.
XX ADQ96506;
XX 07-OCT-2004 (first entry)
XX
XX DE T cell activation associated protein #342.
XX
XX KW antiallergic; antiarthritic; antiaesthatic; antidiabetic; anti-HIV;
XX KW antimicrobial; antirheumatic; immunosuppressive; neuroprotective;
XX KW gene therapy; T cell activation; diagnosis; autoimmune disease;
XX KW rheumatoid arthritis; asthma; multiple sclerosis; diabetes;
XX KW allergic disease; infectious disease; AIDS; chronic rejection; organ;
XX KW bone-marrow transplant.
XX
XX OS Homo sapiens.
XX
XX PN WO2004058805-A2.
XX
XX XX 15-JUL-2004.
XX
XX PD 25-DEC-2003; 2003WO-JP016715.
XX
XX PR 26-DEC-2002; 2002JP-00376365.
XX PR 27-DEC-2002; 2002US-0436473P.
XX PR 25-APR-2003; 2003JP-00122113.
XX PR 28-APR-2003; 2003US-0465792P.
XX PR 21-OCT-2003; 2003JP-00360559.
XX PR 22-OCT-2003; 2003US-0512846P.
XX
XX PA (ASAH-) ASahi Kasei PHARMA CORP.
XX
XX PI Matsuda A, Yoneta S;
XX
XX DR WPI; 2004-593134/57.
XX DR N-PSDB; ADQ96505.
XX
XX PT New purified protein involved in T cell activation, useful for
XX PT diagnosing, preventing and/or treating acquired immunodeficiency
XX PT syndrome, autoimmune (e.g. rheumatoid arthritis, and diabetes), allergic
```

PT	and infectious diseases.	XX		CC	of sites expressing (II). (I) and (II) are useful for treating disorders
PS	Claim 1; SEQ ID NO 684; 2828pp; English.	XX		CC	involving aberrant protein expression or biological activity. The
XX		XX		CC	polypeptide and polynucleotide sequences have applications in
XX		XX		CC	diagnostics, forensics, gene mapping, identification of mutations
CC	The invention relates to purified proteins and genes encoding them, that	CC		CC	responsible for genetic disorders or other traits to assess biodiversity
CC	are involved in T cell activation (I) and has an amino acid deletion,	CC		CC	and to produce other types of data and products dependent on DNA and
CC	substitution or addition in the amino acid sequences. The methods and	CC		CC	amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC	compositions of the present invention are useful for the diagnosis,	CC		CC	amino acid sequences of the invention. Note: The sequence data for this
CC	prevention and/or treatment of autoimmune disease (rheumatoid arthritis,	CC		CC	patent did not appear in the printed specification, but was obtained in
CC	asthma, multiple sclerosis and diabetes), allergic disease, infectious	CC		CC	electronic format directly from WIPO at
CC	disease, AIDS, and acute or chronic rejection at organ transplant or bone	CC		CC	ftp.wipo.int/pub/published_pct_sequences
CC	-marrow transplant. This sequence corresponds to a protein involved in T	XX		XX	Sequence 50 AA;
XX		XX		XX	
SQ	Sequence 561 AA;	XX		XX	
Query Match	86.6%; Score 58; DB 8; Length 561;	Query Match	85.1%; Score 57; DB 4; Length 50;		
Best Local Similarity	80.0%; Pred. No. 1.2e+02;	Best Local Similarity	80.0%; Pred. No. 21;		
Matches	8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;	Matches	8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;		
Qy	1 P P P Y P P P P P I P 10	Qy	1 P P P Y P P P P P I P 10		
Db	201 P P A Y P P P P P V P 210	Db	4 P P P L P P P P P L P 13		
RESULT 37		RESULT 38			
ABG19592		ABB71558			
ID	ABG19592 standard; protein; 50 AA.	ID	ABB71558 standard; protein; 94 AA.		
XX		XX			
AC	ABG19592;	AC	ABB71558;		
XX		XX			
DT	18-FEB-2002 (first entry)	DT	26-MAR-2002 (first entry)		
XX		XX			
DE	Novel human diagnostic protein #19583.	DE	Drosophila melanogaster polypeptide SEQ ID NO 41466.		
XX		XX			
XX	Human; chromosome mapping; gene mapping; gene therapy; forensic;	XX	Drosophila; developmental biology; cell signalling; insecticide;		
KW	food supplement; medical imaging; diagnostic; genetic disorder.	KW	pharmaceutical.		
XX		XX			
OS	Homo sapiens.	OS	Drosophila melanogaster.		
XX		XX			
XX	WO200175067-A2.	XX	WO200171042-A2.		
PN		XX			
XX		XX			
PD	11-OCT-2001.	PD	27-SEP-2001.		
XX		XX			
XX	30-MAR-2001; 2001WO-US008631.	XX	23-MAR-2001; 2001WO-US009231.		
PF		XX			
XX		XX			
PR	31-MAR-2000; 2000US-00540217.	PR	23-MAR-2000; 2000US-0191637P.		
XX		PR	11-JUL-2000; 2000US-00614150.		
XX	23-AUG-2000; 2000US-00649167.	XX			
XX		XX	(PEKE) PE CORP NY.		
PA	(HYSE-) HYSEQ INC.	XX			
XX		XX	Venter JC, Adams M, Li PWD, Myers EW;		
PI	Drmanac RT, Liu C, Tang YT;	XX			
XX		XX			
XX		XX	WPI; 2001-656860/75.		
DR	N-PSDB; AAS83779.	DR	N-PSDB; ABL15661.		
XX		XX			
PT	New isolated polynucleotide and encoded polypeptides, useful in	PT	New isolated nucleic acid detection reagent for detecting 1000 or more		
PT	diagnostics, forensics, gene mapping, identification of mutations	PT	genes from Drosophila and for elucidating cell signaling and cell-cell		
PT	responsible for genetic disorders or other traits and to assess	PT	interactions.		
PT	biodiversity.	PT			
XX		PS	Disclosure; SEQ ID NO 41466; 21pp + Sequence Listing; English.		
XX		XX			
PS	Claim 20; SEQ ID NO 49951; 103pp; English.	XX			
XX		XX			
CC	The invention relates to isolated polynucleotide (I) and polypeptide (II)	CC	The invention relates to an isolated nucleic acid detection reagent		
CC	sequences. (I) is useful as hybridisation probes, polymerase chain	CC	capable of detecting 1000 or more genes from Drosophila. The invention is		
CC	reaction (PCR) primers, oligomers, and for chromosome and gene mapping,	CC	useful in developmental biology and in elucidating cell signalling and		
CC	and in recombinant production of (II). The polynucleotides are also used	CC	cell-cell interactions in higher eukaryotes for the development of		
CC	in diagnostics as expressed sequence tags for identifying expressed	CC	insecticides, therapeutics and pharmaceutical drugs. The invention		
CC	genes. (I) is useful in gene therapy techniques to restore normal	CC	discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA		
CC	activity of (II) or to treat disease states involving (II). (II) is	CC	sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-		
CC	useful for generating antibodies against it, detecting or quantitating a	CC	ABB72072). The sequence data for this patent did not form part of the		
CC	polypeptide in tissue, as molecular weight markers and as a food	CC	printed specification, but was		

```
Query Match      85.1%; Score 57; DB 4; Length 94;
Best Local Similarity 80.0%; Pred. No. 35;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPPPPPPPIP 10
Db 78 PPLPLPLPLP 87

RESULT 39
AAO02495
ID AAO02495 standard; protein; 98 AA.
XX
AC AAO02495;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 16387.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation.
XX
OS Homo sapiens.
XX
PN WO200164835-A2.
XX
DT 07-SEP-2001.
XX
PF 26-FEB-2001; 2001WO-US004927.
XX
PR 28-FEB-2000; 2000US-00515126.
XX
PR 18-MAY-2000; 2000US-00577409.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
WPI; 2001-514838/56.
XX
N-PSDB; AAI82426.
XX
Isolated nucleic acids and polypeptides, useful for preventing diagnosing
PT and treating e.g. leukemia, inflammation and immune disorders.
XX
Claim 20; SEQ ID NO 16387; 1399pp + Sequence Listing; English.
XX
The invention relates to human polynucleotides (AAI79941-AAI93841) and
CC the encoded proteins (AAO0010-AAO13910) that exhibit activity relating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 98 AA;

Query Match      85.1%; Score 57; DB 4; Length 98;
Best Local Similarity 80.0%; Pred. No. 36;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPPPPPPPIP 10
Db 20 PPLPLPLPLP 29

RESULT 40
ABO76127
ID ABO76127 standard; protein; 142 AA.
```

```
ABG16489
ID ABG16489 standard; protein; 118 AA.
XX
AC ABG16489;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #16480.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
DT 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
XX
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
WPI; 2001-639362/73.
XX
N-PSDB; AAS80676.
XX
New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
Claim 20; SEQ ID NO 46848; 103pp; English.
XX
The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABO0010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 118 AA;

Query Match      85.1%; Score 57; DB 4; Length 118;
Best Local Similarity 80.0%; Pred. No. 42;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPPPPPPPIP 10
Db 11 PPLPLPLPLP 20

RESULT 41
ABO76127
ID ABO76127 standard; protein; 142 AA.
```


XX ABO76127;
XX 29-JUL-2004 (first entry)
XX Pseudomonas aeruginosa polypeptide #8302.
XX Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
XX Pseudomonas aeruginosa.
XX US6551795-B1.
XX 22-APR-2003.
XX 18-FEB-1999; 99US-00252991.
XX 18-FEB-1998; 98US-0074788P.
XX 27-JUL-1998; 98US-0094150P.
XX (GENO-) GENOME THERAPEUTICS CORP.
XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;
XX WPI; 2003-615309/58.
XX N-PSDB; ABD09698.
XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
XX useful as molecular targets for diagnostics, prophylaxis and treatment of
XX pathological conditions resulting from bacterial infection.
XX Disclosure; SEQ ID NO 24873; 455pp; English.
XX The invention relates to Pseudomonas aeruginosa polypeptides and the
XX polynucleotides encoding them. The sequences are useful in diagnosis and
XX therapy of pathological conditions, as molecular targets for diagnostics,
XX prophylaxis and treatment of pathological conditions resulting from a
XX bacterial infection, for evaluating a compound, such as a polypeptide,
XX for the ability to bind a P. aeruginosa nucleic acid, as components of
XX effective antibacterial targets, as targets for antibacterial drugs,
XX including anti-P. aeruginosa drugs, as templates for recombinant
XX production of P. aeruginosa-derived peptides or polypeptides, as target
XX components for diagnosis and/or treatment of P. aeruginosa-caused
XX infection, and in detection of P. aeruginosa sequences or other sequences
XX of Pseudomonas species using biotech technology. Sequences ABO67826-
XX ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The
XX sequence data for this patent did not form part of the printed
XX specification but was obtained in electronic format from USPTO at
XX seqdata.uspto.gov/sequence.html
XX
XX Sequence 142 AA;
Query Match 85.1%; Score 57; DB 7; Length 142;
Best Local Similarity 80.0%; Pred. No. 48;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Qy 1 PPPYPPPPPIP 10
Db 89 PPSPPPPPPVP 98
||| |||||:
||| |||||:
RESULT 42
AAB80893
ID AAB80893 standard; protein; 182 AA.
XX AAB80893;
XX 30-MAY-2001 (first entry)
XX Human ATfx leucine zipper domain.
XX Human, ATfx; gamma aminobutyric acid B receptor; GABA B receptor;
XX transcription factor; leucine zipper.

XX Homo sapiens.
XX WO200116596-A2.
XX 08-MAR-2001.
XX 31-AUG-2000; 2000WO-GB003361.
XX 31-AUG-1999; 99GB-00020569.
XX 12-JAN-2000; 2000GE-00000516.
XX (GLAX) GLAXO GROUP LTD.
XX White J, Wise A, Marshall F;
XX WPI; 2001-265904/27.
XX N-PSDB; AAF77878.
XX Identifying modulators of gamma aminobutyric acid(GABA)-B receptor
XX mediated activity by monitoring the interaction between GABAB receptor
XX and the CREB/ATF transcription factors in the presence of a test
XX substance.
XX Claim 14; Page 64-65; 71pp; English.
XX The present invention relates to a method for identifying a modulator of
XX gamma aminobutyric acid B (GABA B) receptor-mediated activity, by
XX monitoring the interaction between a CREB/ATF transcription factor
XX capable of binding to GABA B receptor. The present sequence is the
XX leucine zipper domain of human ATfx (see AAB80892), which was used in the
XX method of the present invention. ATfx is a member of the CREB/ATF family
XX of transcription factors. ATfx contains a bZIP domain, by which it
XX interacts with the coiled coil domain of GABA B receptor. Modulators of
XX GABA B receptor activity are useful for treating central nervous system
XX or peripheral nervous system disorders
XX Sequence 182 AA;
Query Match 85.1%; Score 57; DB 4; Length 182;
Best Local Similarity 80.0%; Pred. No. 59;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Qy 1 PPPYPPPPPIP 10
Db 27 PPELPPEPLP 36
||| |||||:
||| |||||:
RESULT 43
AAB80892
ID AAB80892 standard; protein; 216 AA.
XX AAB80892;
XX 30-MAY-2001 (first entry)
XX Human ATfx.
XX Human, ATfx; gamma aminobutyric acid B receptor; GABA B receptor;
XX transcription factor.
XX Homo sapiens.
XX WO200116596-A2.
XX 08-MAR-2001.
XX 31-AUG-2000; 2000WO-GB003361.
XX 31-AUG-1999; 99GB-00020569.
XX 12-JAN-2000; 2000GB-00000516.
XX (GLAX) GLAXO GROUP LTD.
XX

```
XX
PI White J, Wise A, Marshall F;
DR WPI; 2001-265904/27.
DR N-PSDB; AAF77877.
XX
PT Identifying modulators of gamma aminobutyric acid(GABA)-B receptor
PT mediated activity by monitoring the interaction between GABAB receptor
PT and the CREB/ATF transcription factors in the presence of a test
PT substance.
XX
PS Claim 14; Fig 1; 71pp; English.
XX
CC The present invention relates to a method for identifying a modulator of
CC gamma aminobutyric acid B (GABAB) receptor-mediated activity, by
CC monitoring the interaction between a CREB/ATF transcription factor
CC capable of binding to GABAB receptor. The present sequence is human
CC ATFx, which was used in the method of the present invention. ATFx is a
CC member of the CREB/ATF family of transcription factors. ATFx contains a
CC bZIP domain, by which it interacts with the coiled coil domain of GABAB
CC receptor. Modulators of GABAB receptor activity are useful for treating
CC central nervous system or peripheral nervous system disorders
XX
XX Sequence 216 AA;
SQ
Query Match 85.1%; Score 57; DB 4; Length 216;
Best Local Similarity 80.0%; Pred. No. 68;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 PPYP PPPPIP 10
DB 61 PPPL PPPPLP 70
XX
RESULT 44
ABB50151
ID ABB50151 standard; protein; 216 AA.
XX
AC ABB50151,
XX
DT 05-FEB-2002 (first entry)
XX
DB Human transcription factor TRFX-2.
XX
KW Human; transcription factor; TRFX; cell proliferative disease;
KW autoimmune disease; inflammation; neurological disease;
KW developmental disorder; cancer; AIDS; infection; cycostatic; anti-HIV;
KW neuroprotective; antiinflammatory; gene therapy.
XX
OS Homo sapiens.
XX
PN WO200172777-A2.
XX
PD 04-OCT-2001.
XX
PF 13-MAR-2001; 2001WO-US008117.
XX
PR 13-MAR-2000; 2000US-0188986P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Hillman JL, Baughn MR, Yue H, Lal P, Lu DAM, Patterson C;
PI Azimzai Y, Bandman O, Tang YT, Mathur P, Shah P, Au-Young J;
PI Reddy R;
XX
XX WPI; 2001-570896/64.
DR N-PSDB; ABA82975.
XX
XX Novel transcription factor polypeptides, used to treat diseases
PT associated with altered activity and expression of TRFX, and to screen
PT for agents capable of modulating its activity.
XX
PS Claim 1; Page 140-141; 327pp; English.
XX
XX The present sequence is the protein sequence for a human transcription
XX factor. The transcription factor and its coding sequence are useful in
XX the diagnosis, treatment and prevention of diseases associated with
XX altered expression of the transcription factor e.g. cell proliferative,
XX autoimmune/inflammatory, neurological and developmental disorders. A
XX number of specific disorders/diseases are given in the specification,
XX including: arteriosclerosis, cirrhosis, hepatitis, cancers, AIDS,
XX allergies, anaemia, asthma, autoimmune thyroiditis, bronchitis, atopic
XX dermatitis, diabetes mellitus, emphysema, Goodpasture's syndrome, gout,
XX Grave's disease, multiple sclerosis, osteoarthritis, pancreatitis,
XX psoriasis, rheumatoid arthritis, systemic lupus erythematosus, ulcerative
XX colitis, uveitis, Alzheimer's disease, Huntington's disease, Parkinson's
XX disease, stroke, and viral, bacterial, fungal and protozoal infections
XX
SQ Sequence 216 AA;
Query Match 85.1%; Score 57; DB 4; Length 216;
Best Local Similarity 80.0%; Pred. No. 68;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 PPYP PPPPIP 10
DB 61 PPPL PPPPLP 70
XX
RESULT 45
ABG19594
ID ABG19594 standard; protein; 227 AA.
XX
AC ABG19594;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #19585.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
XX (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS83781.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 20; SEQ ID NO 49953; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
```

CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX SQ Sequence 227 AA;

Query Match 85.1%; Score 57; DB 4; Length 227;
Best Local Similarity 80.0%; Pred. No. 71;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPVPPPPPIP 10
Db 64 PPPLPPLP 73

RESULT 46
AAB80896
ID AAB80896 standard; protein; 282 AA.

XX AC AAB80896;

XX DT 30-MAY-2001 (first entry)

XX DE Human ATfx with N-terminal extension.

XX KW Human; ATfx; gamma aminobutyric acid B receptor; GABA B receptor;
transcription factor.

XX OS Homo sapiens.

XX PN WO200116596-A2.

XX PD 08-MAR-2001.

XX PF 31-AUG-2000; 2000WO-GB0033361.

XX PR 31-AUG-1999; 99GB-00020569.

XX PR 12-JAN-2000; 2000GB-00000516.

XX PA (GLAX) GLAXO GROUP LTD.

XX PI White J, Wise A, Marshall F;

XX WPI; 2001-265904/27.

XX DR N-PSDB; AAF77880.

XX PT Identifying modulators of gamma aminobutyric acid(GABA)-B receptor
mediated activity by monitoring the interaction between GABAB receptor
and the CREB/ATF transcription factors in the presence of a test
substance.

XX PS Claim 14; Fig 9; 71pp; English.

XX CC The present invention relates to a method for identifying a modulator of
gamma aminobutyric acid B (GABA-B) receptor-mediated activity, by
monitoring the interaction between a CREB/ATF transcription factor
capable of binding to GABA-B receptor. The present invention is human
ATfx, which was used in the method of the present invention. ATfx is a
member of the CREB/ATF family of transcription factors. ATfx contains a
bZIP domain, by which it interacts with the coiled coil domain of GABA B
receptor. Modulators of GABA-B receptor activity are useful for treating
central nervous system or peripheral nervous system disorders. The
present sequence has an N-terminal extension compared to AAB80892

XX

SQ Sequence 282 AA;

Query Match 85.1%; Score 57; DB 4; Length 282;
Best Local Similarity 80.0%; Pred. No. 85;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPVPPPPPIP 10
Db 127 PPPLPPLP 136

RESULT 47
AAB95859

ID AAB95859 standard; protein; 282 AA.

XX AC AAB95859;

XX DT 26-JUN-2001 (first entry)

XX DE Human protein sequence SEQ ID NO:18922.

XX KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.

XX OS Homo sapiens.

XX PN EP1074617-A2.

XX PD 07-FEB-2001.

XX PF 28-JUL-2000; 2000EP-00116126.

XX PR 29-JUL-1999; 99JP-00248036.

XX PR 27-AUG-1999; 99JP-00300253.

XX PR 11-JAN-2000; 2000JP-00118776.

XX PR 02-MAY-2000; 2000JP-00183767.

XX PR 09-JUN-2000; 2000JP-00241899.

XX PA (HELI-) HELIX RES INST.

XX PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX WPI; 2001-318749/34.

XX PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-
length cDNAs defined in the specification, and for the detection and/or
diagnosis of the abnormality of the proteins encoded by the full-length
cDNAs.

XX PS Claim 8; SEQ ID NO 18922; 2537pp + Sequence Listing; English.

XX CC The present invention describes primer sets for synthesizing 5602 full-
length cDNAs defined in the specification. Where a primer set comprises:
(a) an oligo-dT primer and an oligonucleotide complementary to the
complementary strand of a polynucleotide which comprises one of the 5602
nucleotide sequences defined in the specification, where the
oligonucleotide comprises at least 15 nucleotides; or (b) a combination
of an oligonucleotide comprising a sequence complementary to the
complementary strand of a polynucleotide which comprises a 5'-end
sequence and an oligonucleotide comprising a sequence complementary to a
polynucleotide which comprises a 3'-end sequence, where the
oligonucleotide comprises at least 15 nucleotides and the combination of
the 5'-end sequence/3'-end sequence is selected from those defined in the
specification. The primer sets can be used in antisense therapy and in
gene therapy. The primers are useful for synthesizing polynucleotides,
particularly full-length cDNAs. The primers are also useful for the
detection and/or diagnosis of the abnormality of the proteins encoded by
the full-length cDNAs. The primers allow obtaining of the full-length
cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
represent human amino acid sequences; and AAH13629 to AAH13632 represent
oligonucleotides, all of which are used in the exemplification of the
present invention

XX SQ Sequence 282 AA;
Query Match 85.1%; Score 57; DB 4; Length 282;
Best Local Similarity 80.0%; Pred. No. 85;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 PPPVPPPPIP 10
||| |||||
Db 127 PPPLPPPLP 136

RESULT 48
ABB81169
ID ABB81169 standard; protein; 282 AA.
XX
AC ABB81169;
XX
DT 25-NOV-2002 (first entry)
XX
DE Human activating transcription factor 5 (ATF5) sequence.
XX
KW Epstein-Barr virus; EBV; TRRAP; cofactor; ATF5; elF1; SUPT6H; human;
KW activating transcription factor 5; translation initiation factor;
KW structure regulator; gene therapy; virucide; cytostatic.
XX
OS Homo sapiens.
XX
PN WO200259606-A2.
XX
PD 01-AUG-2002.
XX
PF 23-JAN-2002; 2002WO-EP000659.
XX
PR 23-JAN-2001; 2001EP-00101452.
XX
PA (GPCB-) GPC BIOTECH AG.
PA (GSFU-) GSF FORSCHUNGSZENTRUM UMWELT & GESUNDHEIT.
XX
PI Kempkes B, Ivanov I, Machl A, Bornkamm G, Santak M;
XX
DR WPI; 2002-643349/69.
DR N-PSDB; ABQ79514.
XX
PT Identifying antagonists or inhibitors for treating or preventing Epstein-Barr virus infections, comprises testing for inhibition or reduction of transcription of a specific gene, and of translation of mRNA transcribed from the gene.
XX
PS Disclosure; Page 34; 45pp; English.
XX
CC The invention relates to identifying antagonists or inhibitors of EBV (Epstein-Barr virus) infection. The method involves testing candidate antagonists or inhibitors for the inhibition or reduction of transcription of a gene (I), and of translation of mRNA transcribed from (I), and determining whether the antagonists or inhibitors, or samples containing them test positive. (I) encodes a polypeptide selected from TRRAP (cofactor), activating transcription factor 5 (ATF5), translation initiation factor (elF1), and structure regulator (SUPT6H), their fragments, derivatives or orthologues. The method is used for identifying antagonists or inhibitors which are useful in treating or preventing EBV infections or other related diseases such as EBV-associated haemophagocytic syndrome, chronic active EBV infection, T-cell lymphoma, natural killer cell leukemia/lymphoma, lymphoproliferative diseases in immunocompromised hosts, Hodgkin's disease, pyothorax-associated B-cell lymphoma, smooth muscle tumours, gastric carcinoma or Burkitt's lymphoma. They are also useful in the preparation of pharmaceutical composition for treating or preventing the EBV infection or related diseases. The polypeptides are useful for identifying the antagonists or inhibitors of EBV infection. Conditional mutants in the genes encoding the polypeptides are useful as markers of potentially EBV infected cells. The present sequence represents the human ATF5 amino acid sequence (GenBank accession no. NM_012068)

XX SQ Sequence 282 AA;
Query Match 85.1%; Score 57; DB 5; Length 282;
Best Local Similarity 80.0%; Pred. No. 85;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 PPPVPPPPIP 10
||| |||||
Db 127 PPPLPPPLP 136

RESULT 49
ADE60923
ID ADE60923 standard; protein; 282 AA.
XX
AC ADE60923;
XX
DT 29-JAN-2004 (first entry)
XX
DE Human Protein Q9Y2D1, SEQ ID NO 6837.
XX
KW Human; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
KW spared nerve injury; SNI; Chung.
XX
OS Homo sapiens.
XX
PN WO2003016475-A2.
XX
PD 27-FEB-2003.
XX
PF 14-AUG-2002; 2002WO-US025765.
XX
PR 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX
PA (GEHO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.
XX
PI Woolf C, D'urso D, Befort K, Costigan M;
XX
DR WPI; 2003-268312/26.
XX
PT New composition comprising two or more isolated polypeptides, useful for preparing a medicament for treating pain in an animal.
XX
PS Claim 1; Page; 1017pp; English.
XX
CC The invention discloses a composition comprising two or more isolated rat or human polynucleotides or a polynucleotide which represents a fragment, derivative or allelic variation of the nucleic acid sequence. Also claimed are a vector comprising the novel polynucleotide, a host cell comprising the vector, a method for identifying a nucleotide sequence which is differentially regulated in an animal subjected to pain and a kit to perform the method, an array, a method for identifying an agent that increases or decreases the expression of the polynucleotide sequence that is differentially expressed in neuronal tissue of a first animal subjected to pain, a method for identifying a compound which regulates the expression of a polynucleotide sequence which is differentially expressed in an animal subjected to pain, a method for identifying a compound that regulates the activity of one or more of the polynucleotides, a method for producing a pharmaceutical composition, a method for identifying a compound or small molecule that regulates the activity in an animal of one or more of the polypeptides given in the specification, a method for identifying a compound useful in treating pain and a pharmaceutical composition comprising the one or more polypeptides or their antibodies. The polynucleotide or the compound that modulates its activity is useful for preparing a medicament for treating pain (e.g. spinal segmental nerve injury (Chung), chronic constriction injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene therapy). The sequence presented is a human protein (shown in Table 2 of

CC the specification) which is differentially expressed during pain. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic form directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 282 AA;
 Query Match 85.1%; Score 57; DB 7; Length 282;
 Best Local Similarity 80.0%; Pred. No. 85;
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPVPPPPIP 10
 ||| |||||
 Db 127 PPPLPPLPLP 136

RESULT 50
 ADP12591
 ID ADP12591 standard; protein; 282 AA.

XX AC ADP12591;

XX DT 12-AUG-2004 (first entry)

XX DE Protein encoded by mRNA of the invention #201.

XX KW transplant rejection; immune system; rheumatoid arthritis; lupus;
 XX KW inflammatory bowel disease; multiple sclerosis; HIV; AIDS.

XX OS Homo sapiens.

XX PN WO2004042346-A2.

XX PD 21-MAY-2004.

XX PF 24-APR-2003; 2003WO-US012946.

XX PR 24-APR-2002; 2002US-00131831.

XX PR 20-DEC-2002; 2002US-00325899.

XX PA (EXPR-) EXPRESSION DIAGNOSTICS INC.

XX PI Wohlgemuth J, Fry K, Woodward R, Ly N, Prentice J, Morris M;
 PI Rosenberg S;

XX DR WPI; 2004-400724/37.

XX PT Diagnosing or monitoring transplant rejection, e.g. heart, kidney, liver,
 PT pancreas, pancreatic islet, lung, bone marrow or stem cell transplant
 PT rejection, in an individual, comprises detecting the expression level of
 PT the genes.

XX PS Claim 65; SEQ ID NO 2600; 1762pp; English.

XX CC The present invention relates to diagnosing or monitoring transplant
 CC rejection, e.g. cardiac or kidney transplant rejection, in an individual
 CC comprises detecting the expression level of one or more genes. The
 CC methods, system and kits are useful in diagnosing or monitoring
 CC transplant rejection, e.g. heart, kidney, liver, pancreas, pancreatic
 CC islet, lung, bone marrow or stem cell transplant rejection,
 CC xenotransplant rejection or mechanical organ replacement rejection, in an
 CC individual. The method is also useful in assessing the immune status of
 CC an individual. The methods are also useful in diagnosing and monitoring
 CC diseases that involve the immune system, e.g. rheumatoid arthritis,
 CC lupus, inflammatory bowel diseases, multiple sclerosis, HIV/AIDS or
 CC viral, bacterial or fungal infection. The present sequence represents a
 CC protein that is encoded by the mRNA of the invention.

XX SQ Sequence 282 AA;

Query Match 85.1%; Score 57; DB 8; Length 282;
 Best Local Similarity 80.0%; Pred. No. 85;
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPVPPPPIP 10
 ||| |||||
 Db 127 PPPLPPLPLP 136

Search completed: April 6, 2006, 09:31:16
 Job time : 168.211 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: April 6, 2006, 09:35:13 ; Search time 21.7895 Seconds
(without alignments)
52.989 Million cell updates/sec

Title: US-10-632-388-312

Perfect score: 67

Sequence: 1 PPPPPPPPIPX 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Database :

1: PIR_80.*

2: PIR1.*

3: PIR3.*

4: PIR4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	58	86.6	559	2	I49444
2	57	85.1	172	2	D41132
3	57	85.1	389	2	S27200
4	57	85.1	820	2	T00645
5	56	83.6	222	2	T43500
6	56	83.6	393	2	JC5614
7	56	83.6	563	2	A87378
8	56	83.6	744	1	Q0BEA7
9	56	83.6	1733	1	B45344
10	56	83.6	1953	2	S63244
11	56	83.6	1958	2	B40505
12	55	82.1	182	2	T30760
13	55	82.1	211	2	S28304
14	55	82.1	281	2	I38707
15	55	82.1	314	2	T48514
16	55	82.1	367	1	S02193
17	55	82.1	433	2	T07910
18	55	82.1	442	2	T34018
19	55	82.1	473	2	B85187
20	55	82.1	520	2	E97813
21	55	82.1	529	2	I38607
22	55	82.1	662	2	D40228
23	55	82.1	681	2	JC5929
24	55	82.1	691	2	T46476
25	55	82.1	798	2	D96563
26	55	82.1	907	2	E96636
27	55	82.1	994	2	S19595
28	55	82.1	1058	2	T13286
29	55	82.1	1064	2	T13963

30	55	82.1	1171	2	T17454
31	55	82.1	1201	2	G86441
32	55	82.1	1206	2	S24407
33	55	82.1	1255	2	T31065
34	55	82.1	1375	2	S48375
35	55	82.1	1468	2	S11515
36	55	82.1	1795	2	F97713
37	55	82.1	2783	1	A41948
38	55	80.6	406	2	T28957
39	54	80.6	710	2	T31502
40	54	80.6	1091	2	T13170
41	53	79.1	160	2	D48232
42	53	79.1	161	2	B48232
43	53	79.1	165	2	C48232
44	53	79.1	196	2	B48232
45	53	79.1	209	2	A48232
46	53	79.1	744	2	E86255
47	53	79.1	1127	2	T32404
48	53	79.1	1258	2	JC5765
49	53	79.1	1315	2	T00381
50	52	77.6	122	2	FQ0452
51	52	77.6	138	2	C96734
52	52	77.6	144	2	E86364
53	52	77.6	149	2	A41132
54	52	77.6	151	2	S10084
55	52	77.6	151	2	JQ1686
56	52	77.6	153	2	S17837
57	52	77.6	186	2	C41132
58	52	77.6	257	2	S32101
59	52	77.6	267	2	T15645
60	52	77.6	269	2	C84707
61	52	77.6	302	2	A96681
62	52	77.6	359	2	T13478
63	52	77.6	378	2	T21651
64	52	77.6	383	2	T39597
65	52	77.6	421	1	S11674
66	52	77.6	428	2	E71415
67	52	77.6	464	2	S22637
68	52	77.6	487	2	S42442
69	52	77.6	494	2	B96534
70	52	77.6	543	2	S25128
71	52	77.6	588	2	T45564
72	52	77.6	645	2	A71416
73	52	77.6	649	2	JN0809
74	52	77.6	651	2	T14763
75	52	77.6	666	2	B70803
76	52	77.6	684	2	T36771
77	52	77.6	757	2	B75437
78	52	77.6	768	2	A87722
79	52	77.6	856	2	D86258
80	52	77.6	941	2	A86404
81	52	77.6	1013	2	T46422
82	52	77.6	1021	2	T23252
83	52	77.6	1132	2	A35098
84	52	77.6	1388	2	T00083
85	52	77.6	1585	2	T31611
86	52	77.6	2123	2	F86348
87	52	77.6	2453	2	S60254
88	52	77.6	2706	2	T28155
89	51	76.1	70	2	A96650
90	51	76.1	134	2	D84672
91	51	76.1	185	2	D75602
92	51	76.1	199	2	S14981
93	51	76.1	221	2	T07176
94	51	76.1	250	2	T16342
95	51	76.1	277	2	H87552
96	51	76.1	278	2	A49266
97	51	76.1	289	2	A43562
98	51	76.1	304	2	H86332
99	51	76.1	306	2	T06435
100	51	76.1	318	2	T29479
101	51	76.1	348	2	A83260
102	51	76.1	391	2	S69192

diaphanous-related
unknown protein [i
formin isoform IV
diaphanous protein
hypothetical prote
formin - mouse
190K antigen precu
alpha-fetoprotein
hypothetical prote
hypothetical prote
diaphanous protein
cysteine-rich exte
cysteine-rich exte
cysteine-rich exte
cysteine-rich exte
cysteine-rich exte
hypothetical prote
hypothetical prote
inositol polyphosp
KRAA0633 protein -
extensin-like prot
hypothetical prote
hypothetical prote
collagen-related p
PHLPSA protein - c
hypothetical prote
hypothetical prote
collagen-related p
hypothetical 16K p
extensin-like prot
sorbin - pig
collagen-related p
PHLPSA protein - c
hypothetical prote
hypothetical prote
unknown protein, 8
hypothetical prote
hypothetical prote
probable inositol
acrosin (EC 3.4.21
probable coll wall
extensin - Volvox
nuclear protein EB
hypothetical prote
61K protein - Auto
hypothetical prote
hypothetical prote
drebrin E (clone 9
hypothetical prote
hypothetical prote
probable integral
ABC transporter, A
protein ZC123.1 [i
protein F501.10 [i
hypothetical prote
hypothetical prote
MHC class III hist
hypothetical prote
hypothetical prote
nuclear receptor c
variant-specific s
protein F2401.6 [i
hypothetical prote
hypothetical prote
extensin class I (
extensin homolog -
hypothetical prote
hypothetical prote
fas ligand - rat
homeotic protein H
T20H2.26 protein -
ribonuclease S5 ho
hypothetical prote
hypothetical membr
serine O-acetylra

103	51	76.1	393	2	PQ0479	piatil extensin-li	176	50	74.6	502	2	A55197	Wiskott-Aldrich sy
104	51	76.1	413	2	H87604	hypothetical prote	177	50	74.6	505	2	S72273	actin-depolymerizi
105	51	76.1	438	2	T12494	hypothetical prote	178	50	74.6	512	2	S28267	phosphorin I prec
106	51	76.1	446	2	T07907	hydroxyproline-ric	179	50	74.6	514	1	A36222	phosphoprotein pho
107	51	76.1	447	2	T49439	hypothetical prote	180	50	74.6	515	2	UT0976	phosphoprotein pho
108	51	76.1	448	2	S33926	Wilms' tumor prote	181	50	74.6	520	2	I84718	RAR-beta1 isoform
109	51	76.1	449	2	A38080	Wilms tumor suscep	182	50	74.6	520	2	G86414	probable protein k
110	51	76.1	449	2	A39692	Wilms' tumor prote	183	50	74.6	524	1	B36222	phosphoprotein pho
111	51	76.1	485	2	A33647	sulfated surface g	184	50	74.6	525	2	A33794	phosphoprotein pho
112	51	76.1	516	2	T00974	probable SF16 prot	185	50	74.6	533	2	S37781	retinoid X recepto
113	51	76.1	530	2	A45690	transactivator EBN	186	50	74.6	555	2	T30349	structural protein
114	51	76.1	571	2	C75530	conserved hypotet	187	50	74.6	558	2	G96522	FltA17.16 [impute
115	51	76.1	599	2	T10798	phosphorin-S - Vo	188	50	74.6	561	2	S17821	protoporphyrin IX
116	51	76.1	610	2	T06690	galactonolactone d	189	50	74.6	561	2	G96752	unknown protein F2
117	51	76.1	715	1	TNBE77	77K alpha trans-in	190	50	74.6	574	2	T05964	probable low-affin
118	51	76.1	760	2	T06291	extensin homolog T	191	50	74.6	581	2	B86408	E3H9.11 protein -
119	51	76.1	814	2	T49207	receptor kinase-li	192	50	74.6	589	2	T05074	hypothetical prote
120	51	76.1	1006	2	G86232	hypothetical prote	193	50	74.6	600	2	T14463	galactonolactone d
121	51	76.1	1248	2	T05059	hypothetical prote	194	50	74.6	602	2	T45760	hypothetical prote
122	51	76.1	1392	2	T51947	probable transcrip	195	50	74.6	622	2	T15467	hypothetical prote
123	51	76.1	1737	2	T00209	MSGF8 protein - hu	196	50	74.6	637	2	T04552	hypothetical prote
124	50.5	75.4	437	2	T14192	extensin homolog T	197	50	74.6	666	2	T31461	probable magnesium
125	50	74.6	141	2	A34043	hypothetical proli	198	50	74.6	678	2	AP1841	Mg chelataase chain
126	50	74.6	142	2	B41132	collagen-related p	199	50	74.6	707	2	S60588	drebrin A - rat
127	50	74.6	148	2	S39206	proline-rich prote	200	50	74.6	742	2	F84643	hypothetical prote
128	50	74.6	161	2	H85120	probable proline-r	201	50	74.6	754	2	A85043	hypothetical prote
129	50	74.6	161	2	F86364	hypothetical prote	202	50	74.6	760	2	B96724	hypothetical prote
130	50	74.6	165	2	C71717	hypothetical prote	203	50	74.6	783	1	A38637	Ras interactor RIN
131	50	74.6	182	2	A44157	spermatophorin Sp2	204	50	74.6	825	2	S54465	YTA12 protein prec
132	50	74.6	195	2	T07735	nodulin-20a - soyb	205	50	74.6	847	2	F96531	hypothetical prote
133	50	74.6	209	2	C89005	protein T24A6.3 [i	206	50	74.6	880	2	T48477	hypothetical prote
134	50	74.6	219	2	T10563	proline-rich prote	207	50	74.6	881	2	T33810	hypothetical prote
135	50	74.6	220	2	S52005	nodulin-30 (NpV30)	208	50	74.6	910	2	A34721	androgen receptor
136	50	74.6	221	2	H82857	TonB protein Xp000	209	50	74.6	911	2	B34721	androgen receptor
137	50	74.6	236	2	T02577	probable AP2 domai	210	50	74.6	919	2	A39248	androgen receptor
138	50	74.6	246	2	T46446	hypothetical prote	211	50	74.6	980	2	S54986	regulatory protein
139	50	74.6	250	1	A31757	homeotic protein H	212	50	74.6	1012	2	T13603	probable N-methyl-
140	50	74.6	251	1	B60492	homeotic protein H	213	50	74.6	1033	2	T33470	hypothetical prote
141	50	74.6	255	2	C36222	phosphoprotein pho	214	50	74.6	1095	2	T20528	hypothetical prote
142	50	74.6	262	2	T06600	acetyl-CoA carboxy	215	50	74.6	1100	2	JC8033	leukocyte formin p
143	50	74.6	294	2	A55477	survival motor neu	216	50	74.6	1110	2	T19673	hypothetical prote
144	50	74.6	301	2	T50001	hypothetical prote	217	50	74.6	1173	2	T31421	C-terminal domain-
145	50	74.6	302	2	C86480	33.2K hypothetical	218	50	74.6	1176	2	T49482	hypothetical prote
146	50	74.6	309	1	S09257	homeotic protein H	219	50	74.6	1213	2	A41724	limb deformity (ld
147	50	74.6	314	2	D71410	hypothetical prote	220	50	74.6	1264	2	A36858	G2R protein - vari
148	50	74.6	328	2	C42756	hypothetical prote	221	50	74.6	1443	2	S05979	steroid hormone re
149	50	74.6	352	2	C96643	hypothetical prote	222	50	74.6	1541	2	T02831	AAA protein L4171.
150	50	74.6	357	2	T47367	hypothetical prote	223	50	74.6	1584	2	T00026	brain-specific ang
151	50	74.6	358	2	A54265	CCAAT/enhancer-bin	224	50	74.6	1607	2	T03022	MAP kinase kinase
152	50	74.6	363	2	T16755	hypothetical prote	225	50	74.6	1621	2	T15264	hypothetical prote
153	50	74.6	365	2	T24955	hypothetical prote	226	50	74.6	1684	2	T02632	hypothetical prote
154	50	74.6	383	2	A82295	hypothetical prote	227	50	74.6	1896	2	B72175	DI5R protein - var
155	50	74.6	388	2	T26922	hypothetical prote	228	50	74.6	1897	2	T28621	hypothetical prote
156	50	74.6	389	2	T03411	protein RFL - rice	229	50	74.6	1952	2	T48814	hypothetical prote
157	50	74.6	395	2	T49575	CCAAT/enhancer bin	230	50	74.6	2649	2	T51023	HD protein - mouse
158	50	74.6	401	2	A48423	engrafted homeodom	231	50	74.6	3119	2	I49729	Huntington disease
159	50	74.6	415	1	A34170	acrosin (EC 3.4.21	232	50	74.6	3144	2	A46068	salivary protein p
160	50	74.6	415	2	T34279	hypothetical prote	233	49	73.1	57	2	S10782	extensin class I (
161	50	74.6	431	2	S47538	acrosin (EC 3.4.21	234	49	73.1	67	2	S14971	extensin class I (
162	50	74.6	432	2	A82222	twitching motility	235	49	73.1	75	2	S14973	proline-rich pepti
163	50	74.6	436	2	B55452	cartilage-derived	236	49	73.1	79	1	FJHUSB	extensin class I (
164	50	74.6	440	2	T49691	glyceroldehyde-3-p	237	49	73.1	90	2	T14972	extensin class I (
165	50	74.6	440	2	S71795	transcription fact	238	49	73.1	117	2	T46322	hypothetical prote
166	50	74.6	445	2	S19930	alpha-amylase (EC	239	49	73.1	132	2	S14970	extensin class I (
167	50	74.6	446	2	A34418	H-2 region II bind	240	49	73.1	183	2	A29356	hydroxyproline-ric
168	50	74.6	448	2	D41727	retinoid X recepto	241	49	73.1	169	2	T07623	extensin homolog H
169	50	74.6	449	2	D87682	OmpA family protei	242	49	73.1	169	2	T34520	hypothetical prote
170	50	74.6	451	2	A41651	retinoic acid rece	243	49	73.1	181	2	S14974	extensin class I (
171	50	74.6	457	2	B88456	protein W03A5.3 [i	244	49	73.1	230	2	T29591	hypothetical prote
172	50	74.6	477	2	JC4386	adenyl cyclase-a	245	49	73.1	242	2	S54156	extensin-like prot
173	50	74.6	483	2	D87752	protein C18E3.6 [i	246	49	73.1	270	2	S74993	hypothetical prote
174	50	74.6	483	2	T15180	hypothetical prote	247	49	73.1	279	2	T02495	hypothetical prote
175	50	74.6	497	2	T27012	hypothetical prote	248	49	73.1	280	2	T11671	extensin-like prot

249	49	73.1	287	2	T34397	hypothetical prote	322	48	71.6	653	2	T47581	SCARECROW1 - Arabi
250	49	73.1	358	2	E87309	hypothetical prote	323	48	71.6	653	2	T51244	SCARECROW protein
251	49	73.1	368	2	C29356	hydroxyproline-ric	324	48	71.6	678	2	B70913	probable penicilli
252	49	73.1	368	2	T51200	hypothetical prote	325	48	71.6	736	2	I51691	dishevelled homolo
253	49	73.1	379	2	S31719	proline-rich prote	326	48	71.6	760	2	F86387	probable pto kinas
254	49	73.1	388	2	S25298	extensin (clone To	327	48	71.6	790	1	F0MVHZ	gag-kit polypeptei
255	49	73.1	440	2	H86312	F2H15.2 protein -	328	48	71.6	1204	2	T19918	hypothetical prote
256	49	73.1	440	2	C7807	Wiskott-Aldrich sy	329	48	71.6	1483	2	E86143	F6F3.12 protein -
257	49	73.1	445	2	S41741	N-acetylmuramoyl-L	330	48	71.6	1607	2	T21982	hypothetical prote
258	49	73.1	445	2	A86113	N-acetylmuramoyl-L	331	48	71.6	1615	2	JE0372	low density lipopr
259	49	73.1	445	2	A31272	N-acetylmuramoyl-L	332	48	71.6	1794	2	T38459	hypothetical diver
260	49	73.1	459	2	T31608	hypothetical prote	333	48	71.6	1841	2	T38091	cell division cont
261	49	73.1	483	2	F71684	vitB10 protein (vi	334	48	71.6	1872	2	T24683	hypothetical prote
262	49	73.1	485	2	T27147	hypothetical prote	335	48	71.6	3848	2	T17414	TipC protein - eli
263	49	73.1	485	2	PC4427	Nck, Ash and phosph	336	47.5	70.9	1409	2	T37188	presynaptic activi
264	49	73.1	489	2	T11622	extensin class 1 p	337	47	70.1	30	2	S59482	hydroxyproline-ric
265	49	73.1	496	2	T48812	hypothetical prote	338	47	70.1	164	2	T26608	hypothetical prote
266	49	73.1	530	2	T50498	myc-like protein -	339	47	70.1	199	2	T07622	probable extensin
267	49	73.1	551	2	S57447	HPBR11-7 protein -	340	47	70.1	212	2	F84862	hypothetical prote
268	49	73.1	577	2	T09024	proline-rich prote	341	47	70.1	215	2	S34163	homeotic protein H
269	49	73.1	583	1	A41129	radixin - mouse	342	47	70.1	217	2	S09754	hypothetical prote
270	49	73.1	583	1	A46127	radixin - human	343	47	70.1	217	2	T15873	hypothetical prote
271	49	73.1	583	1	S39805	radixin - pig	344	47	70.1	219	2	I51382	achaete-scute homo
272	49	73.1	603	2	F96797	probable GT-like t	345	47	70.1	226	2	S41032	hypothetical prote
273	49	73.1	613	2	T00077	gag-like protein -	346	47	70.1	229	1	W4WL8	E4 protein - human
274	49	73.1	639	2	G82919	transcription fact	347	47	70.1	240	2	T25814	hypothetical prote
275	49	73.1	650	2	B87791	protein B0207.1 [i	348	47	70.1	245	1	W4WLS	E4 protein - human
276	49	73.1	697	2	C75525	molybdopterin oxid	349	47	70.1	245	1	W4WLB5	E4 protein - human
277	49	73.1	710	2	D96728	hypothetical prote	350	47	70.1	258	2	S55036	tyrosine-rich hydr
278	49	73.1	734	2	T04876	hypothetical prote	351	47	70.1	264	2	P00478	piatil extensin-li
279	49	73.1	745	2	T51370	hypothetical prote	352	47	70.1	281	2	S65692	E4 protein - human
280	49	73.1	786	2	T01456	extensin homolog F	353	47	70.1	301	2	U01663	hybrid proline-ric
281	49	73.1	917	2	T21870	hypothetical prote	354	47	70.1	303	1	TVHUUD	transforming prote
282	49	73.1	1174	2	T38140	protein-tyrosine-p	355	47	70.1	304	1	W4WL47	E4 protein - human
283	49	73.1	1268	2	T31420	C-terminal domain-	356	47	70.1	320	2	T18319	hypothetical prote
284	49	73.1	1410	2	T47137	hypothetical prote	357	47	70.1	336	2	S29507	metavinculin - pig
285	49	73.1	1560	2	T00080	hypothetical prote	358	47	70.1	345	2	T01348	hypothetical prote
286	49	73.1	1575	2	S88448	synaptojanin, 170K	359	47	70.1	347	2	A43815	transforming prote
287	49	73.1	1802	2	H88444	protein C26E6.12 [360	47	70.1	348	2	D88088	protein B0454.1 [i
288	49	73.1	3005	2	T08841	polyprotein - dour	361	47	70.1	348	2	S52720	homeobox protein g
289	48.5	72.4	225	2	T09964	extensin CYC15 pre	362	47	70.1	411	2	S34164	homeotic protein H
290	48.5	72.4	240	2	T070894	probable pra prote	363	47	70.1	417	2	S47539	homeotic protein H
291	48.5	72.4	249	2	A41497	36K antigen pra	364	47	70.1	432	2	T06782	extensin - soybean
292	48.5	72.4	443	2	G86388	unknown protein [i	365	47	70.1	463	2	T45855	hypothetical prote
293	48.5	72.4	592	2	T070863	hypothetical prote	366	47	70.1	467	2	T34874	hypothetical prote
294	48	71.6	100	2	T17126	hypothetical prote	367	47	70.1	474	2	A46584	adenyl cyclase-a
295	48	71.6	135	2	E96750	hypothetical prote	368	47	70.1	474	2	T10271	capsid-associated
296	48	71.6	185	2	B26669	nodulin-20 precurs	369	47	70.1	508	2	T45867	hypothetical prote
297	48	71.6	207	2	I33154	scleraxis - mouse	370	47	70.1	517	2	E89530	protein H28G03.2 [
298	48	71.6	208	2	A34696	Spec3 protein - se	371	47	70.1	546	2	T02029	DNA-binding protei
299	48	71.6	217	2	T09965	extensin CYC17 pre	372	47	70.1	580	2	T10863	extensin precursor
300	48	71.6	323	2	T49723	hypothetical prote	373	47	70.1	585	2	I58403	H4 protein - human
301	48	71.6	338	1	TWMSPB	transforming prote	374	47	70.1	593	2	S15215	virulence-associat
302	48	71.6	338	2	I53043	transforming prote	375	47	70.1	658	2	T08153	cysteine proteinas
303	48	71.6	350	2	T14191	extensin homolog T	376	47	70.1	707	2	A46302	FTB-associated spl
304	48	71.6	356	1	WJHU2H	homeotic protein H	377	47	70.1	810	2	C70791	probable pona' pro
305	48	71.6	371	1	F0MVCS	gag polyprotein -	378	47	70.1	894	2	F84870	hypothetical prote
306	48	71.6	384	2	H88924	protein F33E11.2 [379	47	70.1	927	2	A48085	transcription fact
307	48	71.6	392	2	B48423	homeotic protein e	380	47	70.1	969	2	T15446	hypothetical prote
308	48	71.6	424	2	A54964	splcosome-associ	381	47	70.1	997	2	T28872	hypothetical prote
309	48	71.6	435	2	A42672	choline kinase (EC	382	47	70.1	1003	2	T34066	hypothetical prote
310	48	71.6	453	2	JK0342	choline kinase (EC	383	47	70.1	1010	2	A33509	coractin-binding
311	48	71.6	474	2	I49572	adenyl cyclase-a	384	47	70.1	1032	2	D83637	vinculin - Caenorh
312	48	71.6	476	2	T46067	hypothetical prote	385	47	70.1	1048	2	T31425	serine/threonine p
313	48	71.6	536	1	F0MVMD	gag polyprotein -	386	47	70.1	1066	2	T10108	C-terminal domain-
314	48	71.6	547	2	C56828	unknown protein F1	387	47	70.1	1100	2	T30967	vinculin - mouse
315	48	71.6	550	2	G70597	probable proteinas	388	47	70.1	1134	2	T30967	transcription acti
316	48	71.6	557	2	A55933	paxillin - human	389	47	70.1	1135	1	A29997	meta-vinculin - hu
317	48	71.6	559	2	B55933	paxillin - chicken	390	47	70.1	1252	2	T14272	hypothetical prote
318	48	71.6	569	2	A46462	T cell activation	391	47	70.1	1311	2	T33757	hypothetical prote
319	48	71.6	575	2	C98346	protein F42G4.3a [392	47	70.1	1400	2	T52359	hypothetical prote
320	48	71.6	594	2	B86456	protein trihelix D	393	47	70.1	1420	2	T37781	probable cytoskele
321	48	71.6	603	2	T22111	hypothetical prote	394	47	70.1	1501	1	B29813	174K ninaC protein

395	47	70.1	1584	2	T18276	protein-tyrosine k	468	46	68.7	502	2	T08776	hypothetical prote
396	47	70.1	3124	2	A40020	collagen alpha 1(X	469	46	68.7	514	2	E83644	chromosomal replic
397	46.5	69.4	306	2	A24354	extensin precursor	470	46	68.7	538	1	F0MV1M	gag polyprotein -
398	46.5	69.4	365	2	T26449	hypothetical prote	471	46	68.7	541	2	T19304	hypothetical prote
399	46.5	69.4	419	2	G70602	hypothetical prote	472	46	68.7	542	2	T29707	hypothetical prote
400	46.5	69.4	457	2	I55976	dihydrolipoamide S	473	46	68.7	544	2	T17547	proline-rich prote
401	46.5	69.4	512	1	F0MVGS	gag polyprotein -	474	46	68.7	551	2	A60047	adenyl cyclase-a
402	46.5	69.4	520	1	F0LJGL	gag polyprotein -	475	46	68.7	554	2	T49833	related to VeA pro
403	46.5	69.4	555	2	T13716	dihydrolipoamide S	476	46	68.7	556	2	D70940	probable PPg prote
404	46.5	69.4	1464	2	T13716	bazooka gene prote	477	46	68.7	558	2	A33616	heterogeneous ribo
405	46.5	69.4	1465	2	T17138	CluAA protein - ra	478	46	68.7	559	2	E44265	ENL (translocation
406	46.5	69.4	1467	2	T18411	latrophilin-1, bra	479	46	68.7	564	2	T45866	hypothetical prote
407	46.5	69.4	1471	2	T17149	CluBA protein - ra	480	46	68.7	581	2	S17150	potassium channel
408	46.5	69.4	1472	2	T18413	latrophilin-1, bra	481	46	68.7	581	2	T34478	hypothetical prote
409	46.5	69.4	1510	2	T17145	CluAB protein - ra	482	46	68.7	612	2	T02414	hypothetical prote
410	46.5	69.4	1515	2	T17156	CluBB protein - ra	483	46	68.7	613	2	A39402	potassium channel
411	46	68.7	80	2	FQ0477	pistil extensin-li	484	46	68.7	624	2	S22703	voltage-gated pota
412	46	68.7	90	2	T18068	hypothetical prote	485	46	68.7	628	2	D86466	69.4K hypothetical
413	46	68.7	91	2	S37486	gene msg3 protein	486	46	68.7	633	2	T34479	hypothetical prote
414	46	68.7	140	2	T19083	hypothetical prote	487	46	68.7	643	2	T27429	hypothetical prote
415	46	68.7	147	2	S37485	gene msg1 protein	488	46	68.7	650	2	S72367	ATP-dependent RNA
416	46	68.7	153	2	T31654	hypothetical prote	489	46	68.7	655	2	AD2422	DNA polymerase III
417	46	68.7	154	2	FQ0476	pistil extensin-li	490	46	68.7	658	2	T04831	probable serine/th
418	46	68.7	156	2	C84556	probable bZIP tran	491	46	68.7	667	2	T17221	hypothetical prote
419	46	68.7	172	2	T27505	hypothetical prote	492	46	68.7	684	2	A56154	Abi substrate ena
420	46	68.7	173	2	T51469	glycine/proline-ri	493	46	68.7	695	2	T24950	hypothetical prote
421	46	68.7	194	2	A38203	proline-rich prote	494	46	68.7	719	2	A88445	protein C2686.10 [
422	46	68.7	204	2	T07679	protein import rec	495	46	68.7	754	2	T06249	protoporphyrin IX
423	46	68.7	222	2	F96711	hypothetical prote	496	46	68.7	758	2	T02525	protoporphyrin IX
424	46	68.7	244	2	S44822	F4482.3 protein -	497	46	68.7	778	2	B86218	protein T2707.20 [
425	46	68.7	248	2	T83558	hypothetical prote	498	46	68.7	798	2	T19864	hypothetical prote
426	46	68.7	250	2	T08908	hypothetical prote	499	46	68.7	798	2	T21369	hypothetical prote
427	46	68.7	253	2	T17312	hypothetical prote	500	46	68.7	814	2	T05537	probable serine/th
428	46	68.7	264	2	B89005	protein T24A6.18 [501	46	68.7	817	2	S51342	verprolin - yeast
429	46	68.7	275	2	T02559	probable spliceoso	502	46	68.7	856	2	T13159	ElB-5SkDa-associat
430	46	68.7	283	2	S13383	hydroxyproline-ric	503	46	68.7	895	2	C86371	99.7K hypothetical
431	46	68.7	322	2	T22403	hypothetical prote	504	46	68.7	899	2	T16204	hypothetical prote
432	46	68.7	325	2	D70728	hypothetical prote	505	46	68.7	921	2	A33718	retinoblastoma pro
433	46	68.7	340	2	S53763	homeotic protein H	506	46	68.7	948	2	A57640	retinoblastoma bin
434	46	68.7	345	2	S12788	transcription fact	507	46	68.7	981	1	F0MVGM	gag-abl polyprotei
435	46	68.7	346	2	T27896	hypothetical prote	508	46	68.7	984	2	A29513	mineralocorticoid
436	46	68.7	369	2	S20500	hydroxyproline-ric	509	46	68.7	1018	2	T43168	hypothetical prote
437	46	68.7	370	2	B88455	protein T15B12.2 [510	46	68.7	1048	2	T23764	hypothetical prote
438	46	68.7	373	2	A70856	probable lpp2 prot	511	46	68.7	1133	2	A54164	sterol regulatory
439	46	68.7	377	2	JC4368	protein farnesyltr	512	46	68.7	1139	2	A49370	E1A-associated cyc
440	46	68.7	377	2	A41625	protein farnesyltr	513	46	68.7	1148	2	T09073	splicing factor S1
441	46	68.7	379	2	T19089	hypothetical prote	514	46	68.7	1217	2	T39427	probable myosin I
442	46	68.7	380	2	S51797	vasodilator-stimul	515	46	68.7	1219	2	S54570	probable membrane
443	46	68.7	384	2	S51796	hypothetical prote	516	46	68.7	1259	2	T16038	hypothetical prote
444	46	68.7	387	2	T21370	hypothetical prote	517	46	68.7	1296	2	T13936	collar protein iso
445	46	68.7	388	2	JC5437	spliceosome-associ	518	46	68.7	1299	1	S06119	membrane protein p
446	46	68.7	393	2	T33103	lin-1 protein - Ca	519	46	68.7	1335	2	T18289	racGAP protein - s
447	46	68.7	407	2	H87341	OmpA family protei	520	46	68.7	1366	2	B84924	hypothetical prote
448	46	68.7	420	2	T46910	hypothetical prote	521	46	68.7	1445	2	A59437	KIAA1204 protein [
449	46	68.7	427	2	T48159	hypothetical prote	522	46	68.7	1461	2	T41643	probable involveme
450	46	68.7	432	2	B96515	hypothetical prote	523	46	68.7	1560	2	T42727	proliferation pote
451	46	68.7	447	2	S37048	cysteine proteinas	524	46	68.7	1663	2	T42092	s-afadin - rat
452	46	68.7	447	2	A84718	hypothetical prote	525	46	68.7	1829	2	T41751	l-afadin - rat
453	46	68.7	450	2	S07051	cysteine proteinas	526	46	68.7	1882	1	GNVYTR	genome polyprotein
454	46	68.7	450	2	S12099	cysteine proteinas	527	46	68.7	2363	2	T38841	probable pre-mRNA
455	46	68.7	461	1	A46077	steroid hormone re	528	46	68.7	2413	2	S34670	aplicing factor PR
456	46	68.7	462	1	A56120	steroid hormone re	529	45.5	67.9	74	2	B40513	hypothetical prote
457	46	68.7	462	2	A40716	steroid hormone re	530	45.5	67.9	135	2	A88951	protein C383.6 [l
458	46	68.7	464	2	JC7143	endoglucanase I -	531	45.5	67.9	198	2	A45067	laminin B1 chain v
459	46	68.7	465	2	A42128	embryonal long ter	532	45.5	67.9	858	2	JG0183	myosin Myok - Dict
460	46	68.7	475	2	A48120	adenyl cyclase-a	533	45.5	67.9	1151	2	T18535	high molecular mas
461	46	68.7	480	2	C86357	hypothetical prote	534	45.5	67.9	1402	2	T17456	cell surface prote
462	46	68.7	487	2	S65310	probable membrane	535	45	67.2	75	2	T26916	hypothetical prote
463	46	68.7	491	2	S67694	probable membrane	536	45	67.2	101	2	T46506	hypothetical prote
464	46	68.7	496	2	S26402	homeotic protein H	537	45	67.2	115	2	T36886	hypothetical prote
465	46	68.7	498	2	T2167	A30L protein - var	538	45	67.2	132	2	H96708	hypothetical prote
466	46	68.7	498	2	C36851	A29L protein - var	539	45	67.2	154	2	T41831	ACMNPV orf91 - Bom
467	46	68.7	498	2	T28571	hypothetical prote	540	45	67.2	190	2	T35570	hypothetical prote

541	45	67.2	194	2	T12504	hypothetical prote	614	44.5	66.4	164	2	T15525	hypothetical prote
542	45	67.2	212	2	S74288	hypothetical prote	615	44.5	66.4	172	2	T23281	hypothetical prote
543	45	67.2	214	2	T10737	extensin-like cell	616	44.5	66.4	316	2	T31880	hypothetical prote
544	45	67.2	214	2	T09854	proline-rich cell	617	44.5	66.4	357	2	PC4293	nuclear factor 1 f
545	45	67.2	220	2	S42879	nodulin-30 - kidn	618	44.5	66.4	461	2	S45568	nuclear factor 1-A
546	45	67.2	269	2	T30468	hypothetical prote	619	44.5	66.4	491	2	T07598	proline-rich prote
547	45	67.2	274	2	T04619	hypothetical prote	620	44.5	66.4	498	2	S45567	nuclear factor 1-A
548	45	67.2	281	2	T29150	hypothetical prote	621	44.5	66.4	509	2	S45565	nuclear factor 1-A
549	45	67.2	301	2	T18788	hypothetical prote	622	44.5	66.4	509	2	JC5428	nuclear factor 1 f
550	45	67.2	313	1	FOVDA	gag polyprotein -	623	44.5	66.4	513	2	T14194	extensin homolog f
551	45	67.2	320	2	AE2842	conserved hypothet	624	44.5	66.4	522	2	S09996	nuclear factor 1-A
552	45	67.2	332	2	S86182	hypothetical prote	625	44.5	66.4	531	2	G85071	hypothetical prote
553	45	67.2	336	2	B47301	Var10 homolog - Bo	626	44.5	66.4	532	2	B36596	nuclear factor 1 -
554	45	67.2	346	2	S76923	hypothetical prote	627	44.5	66.4	688	2	T09941	transcription fact
555	45	67.2	358	2	T01296	leucine-rich repea	628	44.5	66.4	719	2	B86490	F28L22.6 protein -
556	45	67.2	359	2	F97619	hypothetical prote	629	44.5	66.4	950	2	T22592	hypothetical prote
557	45	67.2	367	2	T39172	cyclin-dependent k	630	44.5	66.4	1274	2	T37193	enamelin matrix pr
558	45	67.2	373	2	A47234	homeobox protein H	631	44.5	66.4	2957	2	T33152	hypothetical prote
559	45	67.2	386	2	B72645	hypothetical prote	632	44	65.7	17	2	S59481	hydroxyproline-ric
560	45	67.2	388	2	E72103	hypothetical prote	633	44	65.7	88	4	S00849	hypothetical trans
561	45	67.2	388	2	G81568	hypothetical prote	634	44	65.7	98	2	T02437	hypothetical prote
562	45	67.2	388	2	H86520	hypothetical prote	635	44	65.7	134	2	JC5572	proline-rich prote
563	45	67.2	389	2	B82819	outer membrane pro	636	44	65.7	137	1	SSKV	cupredoxin [valida
564	45	67.2	405	2	T42663	hypothetical prote	637	44	65.7	139	2	T05847	hypothetical prote
565	45	67.2	416	2	S27198	homeotic protein H	638	44	65.7	141	2	T06646	blue copper-bindin
566	45	67.2	416	2	JU0465	extensin precursor	639	44	65.7	155	2	T16855	hypothetical prote
567	45	67.2	427	2	A53798	58K membrane-asso	640	44	65.7	156	1	GNVQL2	genome-linked prot
568	45	67.2	439	2	S51939	chitinase (EC 3.2.	641	44	65.7	156	1	GNVQLL	genome-linked prot
569	45	67.2	440	2	S51614	Algal-CAM - Volvox	642	44	65.7	156	1	GNVQWA	genome-linked prot
570	45	67.2	472	2	T07853	hypothetical prote	643	44	65.7	164	2	H87551	conserved hypothet
571	45	67.2	473	2	T05306	homeobox protein A	644	44	65.7	172	2	A39458	carotene biosynthe
572	45	67.2	474	2	A36240	pyrimidine synthe	645	44	65.7	178	2	S26044	sex-determining pr
573	45	67.2	474	2	S57239	gene forked protei	646	44	65.7	181	2	S53178	core antigen - hep
574	45	67.2	491	2	S76943	hypothetical prote	647	44	65.7	183	2	S53155	core antigen - hep
575	45	67.2	495	1	S31223	transcription fact	648	44	65.7	183	2	S33184	core antigen - hep
576	45	67.2	515	1	T38946	phosphoprotein pho	649	44	65.7	184	2	S26046	sex-determining pr
577	45	67.2	535	2	T45831	Sf16-like protein	650	44	65.7	185	2	S53288	core antigen - hep
578	45	67.2	544	2	S44814	F44B9.3 protein -	651	44	65.7	191	2	B87601	OMP-A family protei
579	45	67.2	578	2	AG01119	probable OMPA-fami	652	44	65.7	197	2	A29648	female-specific tr
580	45	67.2	578	2	AC01779	probable exported	653	44	65.7	198	2	T35752	hypothetical prote
581	45	67.2	604	2	S39885	forked protein - f	654	44	65.7	211	2	B89716	protein F45B8.3 li
582	45	67.2	627	2	T26064	hypothetical prote	655	44	65.7	212	2	S53157	e antigen precurs
583	45	67.2	641	2	S57236	forked protein 2.5	656	44	65.7	218	2	F84748	probable AP2 domai
584	45	67.2	650	2	T22002	hypothetical prote	657	44	65.7	226	2	T35236	hypothetical prote
585	45	67.2	655	1	A55736	RNA-binding protei	658	44	65.7	228	2	S53504	extensin-like prot
586	45	67.2	656	1	A49358	hypothetical prote	659	44	65.7	236	2	T45835	hypothetical prote
587	45	67.2	661	2	T15073	leucine-rich repea	660	44	65.7	241	2	T22216	hypothetical prote
588	45	67.2	684	2	T01267	hypothetical prote	661	44	65.7	279	2	A53062	Fas ligand - mouse
589	45	67.2	696	2	T31538	hypothetical prote	662	44	65.7	281	2	D70845	hypothetical prote
590	45	67.2	699	2	A96529	hypothetical prote	663	44	65.7	283	2	G01926	insulin promoter f
591	45	67.2	708	2	D96711	hypothetical prote	664	44	65.7	291	2	S33209	extensin-like prot
592	45	67.2	711	2	S68443	double-stranded RN	665	44	65.7	299	2	G70784	probable mmp93 pro
593	45	67.2	721	2	E70766	hypothetical prote	666	44	65.7	303	2	JQ1386	hypothetical 33K p
594	45	67.2	725	2	T01268	leucine-rich repea	667	44	65.7	304	2	T48281	hypothetical prote
595	45	67.2	745	2	S13586	triacylglycerol li	668	44	65.7	306	2	T09067	extensin-like prot
596	45	67.2	751	2	T34490	hypothetical prote	669	44	65.7	309	2	B87576	hypothetical prote
597	45	67.2	753	2	A27041	tyrosine kinase-re	670	44	65.7	311	2	T02783	probable homeotic
598	45	67.2	772	2	T13078	KIAA0992 protein -	671	44	65.7	324	2	F84913	probable FCA-relat
599	45	67.2	906	2	A71438	probable resistanc	672	44	65.7	324	2	T24102	hypothetical prote
600	45	67.2	954	1	S68178	mixed-lineage prot	673	44	65.7	331	2	B47236	zinc-finger protei
601	45	67.2	986	2	T33135	hypothetical prote	674	44	65.7	348	2	A49594	enhancer factor pr
602	45	67.2	1076	2	S50536	hypothetical prote	675	44	65.7	348	2	T47494	hypothetical prote
603	45	67.2	1099	2	A56155	tumor suppressor p	676	44	65.7	352	2	T18794	hypothetical prote
604	45	67.2	1110	2	A43253	Large tra-1 protei	677	44	65.7	370	2	JC2204	MAPK-activated pro
605	45	67.2	1196	2	T23832	protein-tyrosine k	678	44	65.7	375	2	T51333	transcription fact
606	45	67.2	1403	2	S24548	homeotic protein p	679	44	65.7	381	2	S52985	cell wall protein
607	45	67.2	1436	2	S57238	forked protein 5.4	680	44	65.7	392	1	F0LQGB	gag polyprotein -
608	45	67.2	1449	2	S57237	forked protein 5.6	681	44	65.7	392	2	S29356	gag protein - bovi
609	45	67.2	2205	1	MNMVRN	nonstructural poly	682	44	65.7	396	2	S39793	MAPK-activated pro
610	45	67.2	2212	2	T28157	erythrocyte membra	683	44	65.7	396	2	A36339	FLO protein - gard
611	45	67.2	2871	2	A55624	fibrillin-1 precur	684	44	65.7	399	2	T49754	homeobox protein -
612	45	67.2	3511	2	A59295	unconventional myo	685	44	65.7	401	2	T24381	hypothetical prote
613	44.5	66.4	139	2	S61885	extensin precursor	686	44	65.7	408	2	G96707	hypothetical prote

687	44	65.7	411	2	A70509	hypothetical prote	760	44	65.7	1460	1	EDBIF	immediate-early pr
688	44	65.7	417	2	G64417	hypothetical prote	761	44	65.7	1612	2	T30805	dutti protein - mo
689	44	65.7	418	2	F82638	hypothetical prote	762	44	65.7	1651	2	T14160	transmembrane rece
690	44	65.7	419	2	T29266	hypothetical prote	763	44	65.7	1784	2	T10532	gag-pol polyprotei
691	44	65.7	425	1	FOVWC	hypothetical prote	764	43.5	64.9	42	2	T07030	extensin - tomato
692	44	65.7	426	2	JQ1696	pist11 extensin-11	765	43.5	64.9	330	2	T07517	probable extensin
693	44	65.7	429	2	T06296	extensin-like prot	766	43.5	64.9	385	2	S78100	MAPK-activated pro
694	44	65.7	433	1	FOLJH2	gag polyprotein -	767	43.5	64.9	578	2	T25647	hypothetical prote
695	44	65.7	434	2	C96515	hypothetical prote	768	43.5	64.9	790	2	F86936	probable integral
696	44	65.7	437	2	S44809	P44B9.8 protein -	769	43.5	64.9	874	2	T15570	hypothetical prote
697	44	65.7	448	2	T06076	proline-rich prote	770	43.5	64.9	913	2	JCS463	alpha-glucosidase
698	44	65.7	449	2	H70652	hypothetical prote	771	43.5	64.9	1250	2	T00454	hypothetical prote
699	44	65.7	451	2	T30603	perlecan homolog 2	772	43.5	64.9	1634	1	JCS500	phosphoinositide 3
700	44	65.7	461	2	T10741	extensin-like prot	773	43.5	64.9	1776	2	G86280	protein T5G21.13 [
701	44	65.7	473	2	B75097	hypothetical prote	774	43	64.2	84	2	T09540	proline rich prote
702	44	65.7	477	2	A47236	zinc-finger protei	775	43	64.2	158	2	A86452	protein F6N18.7 [i
703	44	65.7	478	2	C29514	muscarinic acetyl	776	43	64.2	181	2	JCS233	spermatophorin sp2
704	44	65.7	478	2	F86388	hypothetical prote	777	43	64.2	184	2	S78091	endocuticular prot
705	44	65.7	479	2	S10127	muscarinic acetyl	778	43	64.2	184	2	B95350	protein [imported
706	44	65.7	479	2	S33776	muscarinic acetyl	779	43	64.2	192	2	S76867	hypothetical prote
707	44	65.7	480	2	JH0672	brain factor 1 pro	780	43	64.2	192	2	T30477	hypothetical prote
708	44	65.7	488	2	F86209	protein F22G5.14 [781	43	64.2	197	2	S35252	proline-rich prote
709	44	65.7	494	2	A42170	zinc finger protei	782	43	64.2	212	2	S57330	cathelin-like anti
710	44	65.7	497	2	JCS076	myc-associated zin	783	43	64.2	219	2	D72105	hypothetical prote
711	44	65.7	526	2	A34896	adenylate cyclase-	784	43	64.2	224	2	D72861	AcOrf-91 protein -
712	44	65.7	532	2	JC6170	GATA-transcription	785	43	64.2	226	2	S35435	probable integral
713	44	65.7	532	2	T34235	hypothetical prote	786	43	64.2	228	2	S40463	prophenin (PF-2) p
714	44	65.7	554	2	T02445	probable U4/U6 sma	787	43	64.2	239	2	T03078	conserved hypothe
715	44	65.7	559	2	G84642	hypothetical prote	788	43	64.2	239	2	T36995	hypothetical prote
716	44	65.7	590	1	WMBCB	64K capaid assembl	789	43	64.2	254	2	B84901	hypothetical prote
717	44	65.7	593	2	T24379	hypothetical prote	790	43	64.2	257	2	T10586	small nuclear ribo
718	44	65.7	605	1	Q0B83R	BVRP2 (EC-RP3) pro	791	43	64.2	264	2	T06789	hydroxyproline-ric
719	44	65.7	620	2	S06733	hydroxyproline-ric	792	43	64.2	278	2	S20790	extensin - almond
720	44	65.7	626	2	H71173	hypothetical prote	793	43	64.2	278	2	G86512	hypothetical prote
721	44	65.7	630	2	T02524	probable RING zinc	794	43	64.2	278	2	T10019	hypothetical prote
722	44	65.7	634	2	T00388	hypothetical prote	795	43	64.2	298	2	S53761	triase-phosphate i
723	44	65.7	644	2	S15464	gp70 protein - mur	796	43	64.2	298	2	T18664	hypothetical prote
724	44	65.7	655	2	S40521	FKHR protein - hum	797	43	64.2	307	2	B72677	hypothetical prote
725	44	65.7	664	2	T01368	hypothetical prote	798	43	64.2	312	2	T46255	hypothetical prote
726	44	65.7	670	2	JCS887	signaling mediator	799	43	64.2	322	2	S25299	extensin precursor
727	44	65.7	679	2	S37842	hypothetical prote	800	43	64.2	323	2	S20099	transforming prote
728	44	65.7	694	2	T01005	hypothetical prote	801	43	64.2	329	2	T10064	cytokinin-induced
729	44	65.7	707	2	T14195	extensin homolog T	802	43	64.2	341	1	TVMSJD	transforming prote
730	44	65.7	733	2	T47618	extensin-like prot	803	43	64.2	341	2	JC4051	Jun-D protein - ra
731	44	65.7	756	2	JCS886	signaling mediator	804	43	64.2	346	2	S19129	proline-rich prote
732	44	65.7	789	2	T52067	hypothetical prote	805	43	64.2	355	2	S41285	coat protein - swe
733	44	65.7	808	2	C72858	AcOrf-66 protein -	806	43	64.2	361	2	S19552	potassium channel
734	44	65.7	824	2	T20239	hypothetical prote	807	43	64.2	372	2	T10472	G-box binding prot
735	44	65.7	868	2	T20239	hypothetical prote	808	43	64.2	379	2	A47659	farnesyl-protein t
736	44	65.7	882	2	S41034	hypothetical prote	809	43	64.2	381	2	T27806	hypothetical prote
737	44	65.7	899	2	A35895	androgen receptor	810	43	64.2	383	2	T06753	zinc finger protei
738	44	65.7	902	2	B40494	androgen receptor	811	43	64.2	383	2	FOLJGA	gag polyprotein -
739	44	65.7	928	1	R8HU	retinoblastoma-ass	812	43	64.2	393	2	T43401	transcription init
740	44	65.7	945	2	T00746	hypothetical prote	813	43	64.2	411	2	B41398	inhibin beta-B cha
741	44	65.7	951	2	T47617	extensin-like prot	814	43	64.2	414	2	T50010	hypothetical prote
742	44	65.7	958	2	T13593	hypothetical prote	815	43	64.2	416	2	B75455	N-acetyl-gamma-glu
743	44	65.7	979	2	A35913	regulatory factor	816	43	64.2	418	2	T19800	hypothetical prote
744	44	65.7	1000	2	S44898	ZK1236.3 protein -	817	43	64.2	427	2	I49603	transcription regu
745	44	65.7	1001	2	T28897	hypothetical prote	818	43	64.2	430	2	JC2301	hypothetical 47.8K
746	44	65.7	1017	2	T15598	hypothetical prote	819	43	64.2	444	2	B83802	hypothetical prote
747	44	65.7	1186	2	T42729	histocompatibility	820	43	64.2	445	2	T30604	hypothetical prote
748	44	65.7	1203	2	T51029	related to pathway	821	43	64.2	448	1	A56018	transcription fact
749	44	65.7	1212	2	T42387	histocompatibility	822	43	64.2	449	1	S30205	transcription fact
750	44	65.7	1240	2	JCS209	insulin receptor s	823	43	64.2	450	2	A86919	probable membrane
751	44	65.7	1241	2	T18311	hypothetical prote	824	43	64.2	451	1	A40168	transcription fact
752	44	65.7	1242	2	JS0670	insulin receptor s	825	43	64.2	467	2	S41318	hypothetical prote
753	44	65.7	1262	2	T13353	protein stn-B - fr	826	43	64.2	468	2	S44815	F44B9.4 protein -
754	44	65.7	1306	2	T13592	hypothetical prote	827	43	64.2	469	2	I37451	HBF-2 (HPK-2) pro
755	44	65.7	1308	2	T15280	hypothetical prote	828	43	64.2	476	2	F87324	hypothetical prote
756	44	65.7	1347	2	T02214	ubiquitous TPR mot	829	43	64.2	476	2	A54743	transcription fact
757	44	65.7	1401	2	T02255	probable ubiquitou	830	43	64.2	477	2	I38409	adenylyl cyclase-a
758	44	65.7	1411	2	T48529	hypothetical prote	831	43	64.2	483	2	T25992	hypothetical prote
759	44	65.7	1429	2	T13720	gene expanded prot	832	43	64.2	484	2	T07675	cyclin a2-type, mi

833	43	64.2	486	2	A41537	DNA-binding protei	906	42.5	63.4	176	2	A86441	hypothetical prote
834	43	64.2	487	1	S52261	NADH2 dehydrogenas	907	42.5	63.4	279	2	T05421	hypothetical prote
835	43	64.2	488	2	UC7985	transient receptor	908	42.5	63.4	403	2	S52796	prpL2 protein - hu
836	43	64.2	493	2	T01206	cysteine proteinas	909	42.5	63.4	440	1	S60755	rhodopsin - Allote
837	43	64.2	493	2	T29030	hypothetical prote	910	42.5	63.4	461	2	S34472	MFH-1 protein - mo
838	43	64.2	514	2	T15338	hypothetical prote	911	42.5	63.4	462	2	E70955	hypothetical prote
839	43	64.2	511	2	T16557	hypothetical prote	912	42.5	63.4	517	1	ERADA7	early E2A DNA-bind
840	43	64.2	557	2	S62522	nuclear protein sp	913	42.5	63.4	554	2	F86244	hypothetical prote
841	43	64.2	560	2	T38065	gene NMB protein -	914	42.5	63.4	585	2	S06958	shingomyelin phos
842	43	64.2	560	2	T59302	brain specific Na+	915	42.5	63.4	627	1	S27393	shingomyelin phos
843	43	64.2	574	2	T43556	wiskott-Aldrich sy	916	42.5	63.4	629	1	A39825	shingomyelin phos
844	43	64.2	574	2	T38819	wiskott-Aldrich sy	917	42.5	63.4	679	2	E75262	conserved hypotet
845	43	64.2	575	2	C34106	protein kinase (EC	918	42.5	63.4	1047	2	A55617	masquerade precurs
846	43	64.2	591	1	WNBP02	gene P2 protein -	919	42.5	63.4	1211	2	T42230	AF4 protein - mous
847	43	64.2	591	2	C71460	hypothetical prote	920	42.5	63.4	1217	2	T42625	AF-4 protein - mou
848	43	64.2	592	2	T32402	hypothetical prote	921	42.5	63.4	1297	2	S25714	son-of-sevenless-2
849	43	64.2	614	2	D34106	protein kinase (EC	922	42.5	63.4	1372	2	T43296	cell fusion protei
850	43	64.2	619	2	E84800	hypothetical prote	923	42	62.7	52	2	F71353	hypothetical prote
851	43	64.2	635	1	WNBEW6	capsid protein - h	924	42	62.7	52	2	E98316	hypothetical prote
852	43	64.2	650	2	T04487	hypothetical prote	925	42	62.7	76	2	E97763	protein transport
853	43	64.2	651	2	T46050	hypothetical prote	926	42	62.7	92	2	E48831	vitelline membrane
854	43	64.2	679	2	A42073	potassium channel	927	42	62.7	97	2	E84746	hypothetical prote
855	43	64.2	691	2	F91251	probable tape meas	928	42	62.7	108	2	T26880	hypothetical prote
856	43	64.2	694	2	S68442	Grb2-associated bi	929	42	62.7	118	1	W4W113	E4 protein - human
857	43	64.2	699	2	T05225	extensin homolog F	930	42	62.7	129	2	H75351	conserved hypotet
858	43	64.2	711	2	S43464	ecdysteroid-induce	931	42	62.7	134	2	T36365	proline-rich prote
859	43	64.2	715	2	G86239	protein F20B24.6 {	932	42	62.7	135	2	T49996	ACAGP4 - Arabidops
860	43	64.2	720	2	JO1676	AB13 protein - Ara	933	42	62.7	143	2	S42579	QID3 protein - fun
861	43	64.2	725	2	T00492	hypothetical prote	934	42	62.7	145	2	T48552	glutaredoxin-like
862	43	64.2	727	2	C84534	hypothetical prote	935	42	62.7	149	2	T46937	hypothetical prote
863	43	64.2	733	2	C87655	penicillin-binding	936	42	62.7	159	2	D72486	hypothetical prote
864	43	64.2	769	2	T56546	Shaw type potasinu	937	42	62.7	160	2	S58759	ezrin - rat (fragm
865	43	64.2	790	2	T25045	hypothetical prote	938	42	62.7	165	2	T24470	hypothetical prote
866	43	64.2	793	1	S60735	splicing factor SF	939	42	62.7	168	2	T48343	hypothetical prote
867	43	64.2	872	2	S33015	hypothetical prote	940	42	62.7	168	2	T48375	transcription co-a
868	43	64.2	872	2	T50369	probable serine/th	941	42	62.7	173	2	T19341	hypothetical prote
869	43	64.2	876	2	T49801	hypothetical prote	942	42	62.7	178	2	E87104	conserved hypotet
870	43	64.2	892	2	T09071	SH3 domains-contai	943	42	62.7	185	2	A82752	hypothetical prote
871	43	64.2	909	2	T06635	hypothetical prote	944	42	62.7	191	2	F84522	probable proline-r
872	43	64.2	924	2	T06636	hypothetical prote	945	42	62.7	191	2	D72711	hypothetical prote
873	43	64.2	934	2	T08418	protein kinase (EC	946	42	62.7	192	2	T38535	probable transloca
874	43	64.2	971	2	T32883	hypothetical prote	947	42	62.7	194	2	AE2310	hypothetical prote
875	43	64.2	992	2	A31666	hypothetical prote	948	42	62.7	199	2	S23635	gamma-coixin, 22K,
876	43	64.2	1038	2	JO7663	ras GTPase-activat	949	42	62.7	211	2	S55129	transcription fact
877	43	64.2	1046	2	T26384	hypothetical prote	950	42	62.7	224	2	C48652	transfer protein s
878	43	64.2	1069	2	S27922	nuclear antigen EB	951	42	62.7	265	2	T46089	proline-rich prote
879	43	64.2	1086	2	T33893	hypothetical prote	952	42	62.7	268	2	H84684	En/Spm-like transp
880	43	64.2	1098	2	T08599	probable transcrip	953	42	62.7	270	2	S63049	hypothetical prote
881	43	64.2	1135	2	T30531	Scythe protein - A	954	42	62.7	287	2	T05338	hypothetical prote
882	43	64.2	1171	2	T13085	PIP82 protein - fr	955	42	62.7	293	2	G70896	hypothetical prote
883	43	64.2	1188	2	S49915	extensin-like prot	956	42	62.7	306	2	T52340	cell wall-plasma m
884	43	64.2	1201	2	A57369	anillin - fruit fl	957	42	62.7	307	2	T22128	hypothetical prote
885	43	64.2	1249	2	A56511	myosin I myoA - Em	958	42	62.7	309	2	S08343	hypothetical prote
886	43	64.2	1323	2	S27224	N-methyl-D-asparta	959	42	62.7	311	2	H70911	hypothetical prote
887	43	64.2	1323	2	T78557	N-methyl-D-asparta	960	42	62.7	318	2	S33158	hypothetical prote
888	43	64.2	1356	1	C45219	N-methyl-D-asparta	961	42	62.7	326	2	A46676	extensin - common
889	43	64.2	1364	2	T00250	MEGF2 protein - hu	962	42	62.7	330	2	S22140	CD68 homolog macro
890	43	64.2	1371	2	T29019	hypothetical prote	963	42	62.7	342	2	A96511	nodulin Enod2 - Se
891	43	64.2	1426	2	T30817	homeotic protein C	964	42	62.7	358	2	JC4311	CCAAT/enhancer bin
892	43	64.2	1557	2	T13160	protein CNK - frui	965	42	62.7	365	2	B42832	factor VIII intron
893	43	64.2	1655	2	T32633	hypothetical prote	966	42	62.7	367	2	F87340	Rieske 2Fe-2S fami
894	43	64.2	1715	2	T13049	eyelid - fruit fly	967	42	62.7	375	2	T51854	RING-H2 finger pro
895	43	64.2	2810	2	T22298	hypothetical prote	968	42	62.7	382	2	E85082	hypothetical prote
896	43	64.2	3002	2	A47221	fibrillin 1 precur	969	42	62.7	382	2	T14186	hypothetical prote
897	43	64.2	3229	2	S27852	probable cell-surf	970	42	62.7	393	2	T05532	hypothetical prote
898	43	64.2	4056	2	H96599	protein F14U16.10	971	42	62.7	394	2	T33641	hypothetical prote
899	43	64.2	4957	2	T03455	ALR protein - huma	972	42	62.7	407	1	A40150	inhibin beta-B cha
900	43	64.2	5262	2	T03454	ALR protein - huma	973	42	62.7	408	2	T34467	hypothetical prote
901	43	64.2	7962	2	T38346	elastic titin - hu	974	42	62.7	413	2	T03240	F10/19F protein ho
902	43	64.2	13288	2	T30399	mucin, submaxillar	975	42	62.7	414	2	JN0866	nucleolar protein
903	42.5	63.4	134	2	A95995	hypothetical prote	976	42	62.7	416	1	FOCH	transforming prote
904	42.5	63.4	150	2	C72687	hypothetical prote	977	42	62.7	416	1	TVEFAC	transforming prote
905	42.5	63.4	162	2	C72665	hypothetical prote	978	42	62.7	421	2	T00676	hypothetical prote

979 42 62.7 423 1 TVFV28 transforming prote
980 42 62.7 427 2 B96804 nucellin-like prot
981 42 62.7 429 2 S77528 hypothetical prote
982 42 62.7 440 2 S65358 familial Alzheimer
983 42 62.7 445 2 A60488 histidine-rich gly
984 42 62.7 448 2 T45710 H-protein promoter
985 42 62.7 451 1 TVFV2C gag-myc polyprotei
986 42 62.7 455 1 OOOCC rhodopsin - giant
987 42 62.7 456 2 S23104 choline kinase - h
988 42 62.7 467 2 S71169 protein kinase, 54
989 42 62.7 478 2 JC4940 synapsin IIb - hum
990 42 62.7 481 2 B86433 protein T17H7.3 [i
991 42 62.7 482 2 E97748 virB10 protein [im
992 42 62.7 483 2 A84698 probable RNA-bind
993 42 62.7 489 2 T26069 hypothetical prote
994 42 62.7 490 2 A35312 potassium channel
995 42 62.7 495 2 T27936 hypothetical prote
996 42 62.7 503 1 LU8011 annexin XI form A
997 42 62.7 505 1 S23447 annexin XI - huma
998 42 62.7 505 2 A53152 hypothetical prote
999 42 62.7 506 2 AG1864 hypothetical prote
1000 42 62.7 548 2 T14793

ALIGNMENTS

RESULT 1
I49444
SH3 binding protein - mouse
C/Species: Mus musculus (house mouse)
C/Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
C/Accession: I49444
R/Ren, R.; Mayer, B.J.; Cicchetti, P.; Baltimore, D.
Science 259, 1157-1161, 1993
A/Title: Identification of a ten-amino acid proline-rich SH3 binding site.
A/Reference number: I49444; MUID:93174278; PMID:8438166
A/Accession: I49444
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-559 <RES>
A/Cross-references: UNIPROT:Q06649; UNIPARC:UPI0000027A0E; GB:L14543; NID:g293267; PIDN:
F/25-128/Domain: pleckstrin repeat homology <PLK>
F/201-210/Region: proline-rich SH3 binding

Query Match 86.68; Score 58; DB 2; Length 559;
Best Local Similarity 80.0%; Pred. No. 8.7;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPPYPPPPIP 10
||| |||||
DB 201 PPAYPPPPVP 210

RESULT 2
D41132
collagen-related protein 4 - Hydra magnipapillata (fragment)
C/Species: Hydra magnipapillata
C/Date: 05-Jun-1992 #sequence_revision 12-Jun-1992 #text_change 15-Sep-2003
C/Accession: D41132; S21932
R/Kurz, E.M.; Holstein, T.W.; Petri, B.M.; Engel, J.; David, C.N.
J. Cell Biol. 115, 1159-1169, 1991
A/Title: Mini-collagens in hydra nematocytes.
A/Reference number: A41132; MUID:92064646; PMID:1955459
A/Accession: D41132
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-172 <KUR>
A/Cross-references: UNIPARC:UPI0000076DA2; EMBL:X61048; NID:g9452; PIDN:CAA43382.1; PID:
A/Note: the authors translated the codon GGT for residue 142 as Pro, and TTA for residue
A/Note: submitted to the EMBL Data Library, July 1991

Query Match 85.1%; Score 57; DB 2; Length 172;
Best Local Similarity 90.0%; Pred. No. 8.7;
Matches 9; Conservative 1; Mismatches 1; Indels 1; Gaps 0;

QY 1 PPPYPPPPIP 10
||| |||||
DB 201 PPAYPPPPVP 210

Best Local Similarity 90.0%; Pred. No. 3.4;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPPYPPPPIP 10
||| |||||
DB 57 PPPPPPPPIP 66

RESULT 3
S27200
proline-rich protein - mouse
C/Species: Mus musculus (house mouse)
C/Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
C/Accession: S27200
R/Sazuka, T.; Tomooka, Y.; Kathju, S.; Ikawa, Y.; Noda, M.; Kumar, S.
Biochim. Biophys. Acta 1132, 240-248, 1992
A/Title: Identification of a developmentally regulated gene in the mouse central nervous
A/Reference number: S27200; MUID:93041923; PMID:1420303
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-389 <SAZ>
A/Cross-references: UNIPROT:Q03173; UNIPARC:UPI000002954A; GB:D10727; NID:g220499; PIDN:

Query Match 85.1%; Score 57; DB 2; Length 389;
Best Local Similarity 80.0%; Pred. No. 7.8;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPPYPPPPIP 10
||| |||||
DB 42 PPPLPPPPLP 51

RESULT 4
T00645
hypothetical protein F3I6.8 - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 09-Jul-2004
C/Accession: T00645
R/Federspiel, N.A.; Palm, C.J.; Conway, A.B.; Kurtz, D.B.; Conway, A.R.; Au, M.; Araujo,
; Vysotskaia, V.S.; Yu, G.; Ecker, J.; Theologis, A.; Davis, R.W.
submitted to the EMBL Data Library, February 1998
A/Reference number: Z14197
A/Accession: T00645
A/Status: translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-820 <FED>
A/Cross-references: UNIPROT:O48682; UNIPARC:UPI00000A0F94; EMBL:AC002396; NID:g2749918;
C/Genetics:
A/Gene: ATSP:F3I6.8
A/Map position: 1
A/Introns: 335/2; 450/3; 676/2

Query Match 85.1%; Score 57; DB 2; Length 820;
Best Local Similarity 90.0%; Pred. No. 17;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPPYPPPPIP 10
||| |||||
DB 238 PPPPPPPPIP 247

RESULT 5
T43500
hypothetical protein DKFZp586G1721.1 - human (fragment)
C/Species: Homo sapiens (man)
C/Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jan-2000
C/Accession: T43500
R/Oftenwaelder, B.; Obermaier, B.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, December 1999
A/Reference number: Z22515
A/Accession: T43500
A/Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-222 <AAA>

A:Cross-references: UNIPARC:UPI000016AC9F; EMBL:AL133642

A:Experimental source: adult uterus; clone DKFZp586G1721

C:Genetics:

A:Note: DKFZp586G1721.1

Query Match 83.6%; Score 56; DB 2; Length 222;
Best Local Similarity 80.0%; Pred. No. 5.7;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPYPPPPPIP 10
||| ||||:|
Db 20 PPPPPPPPPVP 29

RESULT 6

JC5614

RNB6 protein - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 23-Sep-1997 #sequence_revision 17-Oct-1997 #text_change 09-Jul-2004

C:Accession: JC5614

R;Ohta, S.; Mineta, T.; Kimoto, M.; Tabuchi, K.

Biochem. Biophys. Res. Commun. 237, 307-312, 1997

A:Title: Differential display cloning of a novel rat cDNA (RNB6) that shows high expression

A:Reference number: JC5614; MUID:97415794; PMID:9288706

A:Accession: JC5614

A:Molecule type: mRNA

A:Residues: 1-393 <OHT>

A:Cross-references: UNIPROT:O08719; UNIPARC:UPI000012A2AB; GB:U70211; NID:g2058461; PIDN

A:Experimental source: brain

C:Comment: This protein belongs to Ena/VASP family member, and is involved in the development.

Query Match 83.6%; Score 56; DB 2; Length 393;
Best Local Similarity 80.0%; Pred. No. 10;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPYPPPPPIP 10
||| ||||:|
Db 180 PPPPPPPPPVP 189

RESULT 7

A87378

hypothetical protein CC1037 [imported] - Caulobacter crescentus

C:Species: Caulobacter crescentus

C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004

C:Accession: A87378

R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.

B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon

n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.

Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A:Title: Complete Genome Sequence of Caulobacter crescentus.

A:Reference number: A87249; MUID:21173698; PMID:11259647

A:Accession: A87378

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-563 <STO>

A:Cross-references: UNIPROT:Q9A9F1; UNIPARC:UPI00000C7254; GB:AE005673; NID:gl3422331; F

C:Genetics:

A:Gene: CC1037

Query Match 83.6%; Score 56; DB 2; Length 563;
Best Local Similarity 80.0%; Pred. No. 15;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPYPPPPPIP 10
||| ||||:|
Db 322 PPPPPPPPPVP 331

RESULT 8

A87378

hypothetical protein CC1037 [imported] - Caulobacter crescentus

C:Species: Caulobacter crescentus

C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004

C:Accession: A87378

R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.

QOBEA7

UL69 protein - human cytomegalovirus (strain AD169)

C:Species: human cytomegalovirus, human herpesvirus 5

A:Note: host Homo sapiens (man)

C:Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 09-Jul-2004

C:Accession: S09832

R;Chee, M.S.; Bankier, A.T.; Beck, S.; Bohni, R.; Brown, C.M.; Cerny, R.; Horsnell, T.; M.; Barrell, B.G.

Curr. Top. Microbiol. Immunol. 154, 125-169, 1990

A:Title: Analysis of the protein-coding content of the sequence of human cytomegalovirus

A:Reference number: S09749; MUID:90269039; PMID:2161319

A:Accession: S09832

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-744 <CHE>

A:Cross-references: UNIPROT:P16749; UNIPARC:UPI000004735B; EMBL:X17403; NID:g59591; PIDN

A:Note: possible protein-coding frames are given

A:Note: the DNA sequence was submitted to EMBL, December 1989, in computer-readable form

C:Superfamily: cytomegalovirus UL69 protein

C:Keywords: transcription regulation

Query Match 83.6%; Score 56; DB 1; Length 744;
Best Local Similarity 80.0%; Pred. No. 19;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPYPPPPPIP 10
||| ||||:|
Db 708 PPPPPPPPPVP 717

RESULT 9

B45344

probable nuclear antigen - suid herpesvirus 1 (strain Kaplan)

C:Species: suid herpesvirus 1

C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004

C:Accession: B45344

R;Vlcek, C.; Kozmik, Z.; Paces, V.; Schirm, S.; Schwyzer, M.

Virology 179, 363-377, 1990

A:Title: Pseudorabies virus immediate-early gene overlaps with an oppositely oriented op

A:Reference number: A45344; MUID:91021039; PMID:2171211

A:Accession: B45344

A:Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-1733 <VLC>

A:Cross-references: UNIPROT:P33485; UNIPARC:UPI0000138ADE; GB:M34651; NID:g334070; PIDN:

C:Superfamily: pseudorabies virus 1 nuclear antigen

Query Match 83.6%; Score 56; DB 1; Length 1733;
Best Local Similarity 80.0%; Pred. No. 46;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPYPPPPPIP 10
||| ||||:|
Db 276 PPPPPPPPLP 285

RESULT 10

S63244

BN11 protein - Yeast (Saccharomyces cerevisiae)

N;Alternate names: protein N0646; protein YNL271c

C:Species: Saccharomyces cerevisiae

C:Date: 27-Apr-1996 #sequence_revision 03-May-1996 #text_change 09-Jul-2004

C:Accession: S63244; S63245; S48523; S60909; S65111

R;Sen-Gupta, M.; Lyck, R.; Niedenthal, R.K.; Fleig, U.N.; Hegemann, J.H.

submitted to the Protein Sequence Database, April 1996

A:Reference number: S63235

A:Accession: S63244

A:Molecule type: DNA

A:Residues: 1-1553 <SEN>

A:Cross-references: UNIPROT:P41832; UNIPARC:UPI0000168B2A; EMBL:Z71547; MIPS:YNL271c

A:Experimental source: strain S288C

R;Messenguy, F.; Dubois, E.; Vierendeels, F.; Scherens, B.; Pierard, A.; Glansdorff, N.

submitted to the Protein Sequence Database, April 1996

A;Reference number: S63245
A;Accession: S63245
A;Molecule type: DNA
A;Residues: 997-1953 <MBS>
A;Cross-references: UNIPARC:UPI000017B262; EMBL:Z71547; MIPS:YNL271C
A;Experimental source: strain S28C
R;Fares, H.F.; Pringle, J.R.
Submitted to the EMBL Data Library, April 1994
A;Description: Synthetic Lethals of CDC12.
A;Reference number: S48524
A;Accession: S48523
A;Molecule type: DNA
A;Residues: 1-937, 'A', 939-1429, 'C', 1431-1953 <PAR>
A;Cross-references: UNIPARC:UPI0000168E7F; EMBL:L31766; NID:G472524; PIDN:AAA34455.1; PI
R;Sen-Gupta, M.; Lyck, R.; Fleig, U.; Niedenthal, R.N.; Hegemann, J.M.
Submitted to the EMBL Data Library, October 1995
A;Description: The sequence of a 24152 bp segment from the left arm of chromosome XIV fr
A;Reference number: S60909
A;Accession: S60909
A;Molecule type: DNA
A;Residues: 1-1553 <SE2>
A;Cross-references: UNIPARC:UPI0000168B2A; EMBL:X92494; NID:gl045236; PIDN:CAA63225.1; H
R;Sen-Gupta, M.; Lyck, R.; Fleig, U.; Niedenthal, R.K.; Hegemann, J.H.
Yeast 12, 505-514, 1996
A;Title: The sequence of a 24 152 bp segment from the left arm of chromosome XIV from Sa
A;Reference number: S65111; MUID:96310631; PMID:8740425
A;Accession: S65111
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-1553 <SEW>
A;Cross-references: UNIPARC:UPI0000168B2A; EMBL:X92494; NID:gl045236; PIDN:CAA63225.1; H
A;Note: The nucleotide sequence was submitted to the EMBL Data Library, October 1995
C;Genetics:
A;Gene: SGD:BN11; SHE5; SYL39
A;Cross-references: SGD:S0005215; MIPS:YNL271C
A;Map position: 14L

Query Match 83.6%; Score 56; DB 2; Length 1953;
Best Local Similarity 80.0%; Pred. No. 51;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPPPPPPIP 10
||| |||||:
Db 1241 PPPPPPPVP 1250

RESULT 11
B40505
hypothetical protein - suid herpesvirus 1 (strain Indiana-Funkhuser or Becker)
C;Species: suid herpesvirus 1
C;Date: 10-Apr-1992 #sequence_revision 10-Apr-1992 #text_change 09-Jul-2004
C;Accession: B40505
R;Cheung, A.K.
J. Virol. 65, 5260-5271, 1991
A;Title: Cloning of the latency gene and the early protein 0 gene of pseudorabies virus.
A;Reference number: A40505; MUID:91374576; PMID:1654441
A;Accession: B40505
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1958 <CHE>
A;Cross-references: UNIPROT:Q69340; UNIPARC:UPI00000F2308; GB:M57505; NID:G334066; PIDN:
C;Superfamily: pseudorabies virus 1 nuclear antigen

Query Match 83.6%; Score 56; DB 2; Length 1958;
Best Local Similarity 80.0%; Pred. No. 51;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPPPPPPIP 10
||| |||||:
Db 485 PPPPPPPPLP 494

RESULT 12
B40505
hypothetical protein - suid herpesvirus 1 (strain Indiana-Funkhuser or Becker)
C;Species: suid herpesvirus 1
C;Date: 10-Apr-1992 #sequence_revision 10-Apr-1992 #text_change 09-Jul-2004
C;Accession: B40505
R;Cheung, A.K.
J. Virol. 65, 5260-5271, 1991
A;Title: Cloning of the latency gene and the early protein 0 gene of pseudorabies virus.
A;Reference number: A40505; MUID:91374576; PMID:1654441
A;Accession: B40505
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1958 <CHE>
A;Cross-references: UNIPROT:Q69340; UNIPARC:UPI00000F2308; GB:M57505; NID:G334066; PIDN:
C;Superfamily: pseudorabies virus 1 nuclear antigen

Query Match 83.6%; Score 56; DB 2; Length 1958;
Best Local Similarity 80.0%; Pred. No. 51;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPPPPPPIP 10
||| |||||:
Db 485 PPPPPPPPLP 494

RESULT 13
B28304
hypothetical protein T3G5.2 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 12-Mar-1993 #sequence_revision 12-Mar-1993 #text_change 09-Jul-2004
C;Accession: S28304
R;Berks, M.
Submitted to the EMBL Data Library, December 1992
A;Reference number: S28296
A;Accession: S28304
A;Molecule type: DNA
A;Residues: 1-211 <BER>
A;Cross-references: UNIPROT:Q03607; UNIPARC:UPI000013BA56; EMBL:Z19158; NID:G6880; PID:G
C;Genetics:
A;Introns: 138/2; 176/3

Query Match 82.1%; Score 55; DB 2; Length 211;
Best Local Similarity 80.0%; Pred. No. 7;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPPPPPPIP 10
||| |||||:
Db 104 PPPPPPPPLP 113

RESULT 14
I38707
Fas ligand - human
C;Species: Homo sapiens (man)
C;Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 09-Jul-2004
C;Accession: I38707; JC2340; S57565; I38554
R;Takahashi, T.; Tanaka, M.; Inazawa, J.; Abe, T.; Suda, T.; Nagata, S.
Int. Immunol. 6, 1567-1574, 1994
A;Title: Human Fas ligand: gene structure, chromosomal location and species specificity.
A;Reference number: I38707; MUID:95127560; PMID:7826947
A;Accession: I38707
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-281 <RES>
A;Cross-references: UNIPROT:P48023; UNIPARC:UPI00000D91A; EMBL:U11821; NID:G595430; PID
R;Mita, E.; Hayashi, N.; Iio, S.; Takehara, T.; Hijioaka, T.; Kasahara, A.; Fusamoto, H.;
Biochem. Biophys. Res. Commun. 204, 468-474, 1994
A;Title: Role of Fas ligand in apoptosis induced by hepatitis C virus infection.
A;Reference number: JC2340; MUID:95071350; PMID:7980502
A;Accession: JC2340
A;Molecule type: DNA

T30760
hypothetical protein 158R - Molluscum contagiosum virus 1
N;Alternate names: MC158R
C;Species: Molluscum contagiosum virus 1
C;Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 09-Jul-2004
C;Accession: T30760
R;Senkevich, T.G.; Bugert, J.J.; Sisler, J.R.; Koonin, E.V.; Darai, G.; Moss, B.
Science 273, 813-816, 1996
A;Title: Genome sequence of a human tumorigenic poxvirus: Prediction of specific host re
A;Reference number: Z20876; MUID:96325459; PMID:8670425
A;Accession: T30760
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-182 <SEN>
A;Cross-references: UNIPROT:Q98324; UNIPARC:UPI00000EBB89; EMBL:U60315; PIDN:AAC55286.1
C;Genetics:
A;Note: MC158R

Query Match 82.1%; Score 55; DB 2; Length 182;
Best Local Similarity 80.0%; Pred. No. 6.1;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPPPPPPIP 10
||| |||||:
Db 104 PPPPPPPPLP 113

RESULT 13
B28304
hypothetical protein T3G5.2 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 12-Mar-1993 #sequence_revision 12-Mar-1993 #text_change 09-Jul-2004
C;Accession: S28304
R;Berks, M.
Submitted to the EMBL Data Library, December 1992
A;Reference number: S28296
A;Accession: S28304
A;Molecule type: DNA
A;Residues: 1-211 <BER>
A;Cross-references: UNIPROT:Q03607; UNIPARC:UPI000013BA56; EMBL:Z19158; NID:G6880; PID:G
C;Genetics:
A;Introns: 138/2; 176/3

Query Match 82.1%; Score 55; DB 2; Length 211;
Best Local Similarity 80.0%; Pred. No. 7;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPPPPPPIP 10
||| |||||:
Db 90 PPPPPPPPLP 99

RESULT 14
I38707
Fas ligand - human
C;Species: Homo sapiens (man)
C;Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 09-Jul-2004
C;Accession: I38707; JC2340; S57565; I38554
R;Takahashi, T.; Tanaka, M.; Inazawa, J.; Abe, T.; Suda, T.; Nagata, S.
Int. Immunol. 6, 1567-1574, 1994
A;Title: Human Fas ligand: gene structure, chromosomal location and species specificity.
A;Reference number: I38707; MUID:95127560; PMID:7826947
A;Accession: I38707
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-281 <RES>
A;Cross-references: UNIPROT:P48023; UNIPARC:UPI00000D91A; EMBL:U11821; NID:G595430; PID
R;Mita, E.; Hayashi, N.; Iio, S.; Takehara, T.; Hijioaka, T.; Kasahara, A.; Fusamoto, H.;
Biochem. Biophys. Res. Commun. 204, 468-474, 1994
A;Title: Role of Fas ligand in apoptosis induced by hepatitis C virus infection.
A;Reference number: JC2340; MUID:95071350; PMID:7980502
A;Accession: JC2340
A;Molecule type: DNA

A;Residues: 1-281 <MIT>
A;Cross-references: UNIPARC:UPI000000D91A; GB:D38122; DDBJ:D29820; NID:G601892; PIDN:BA
R;Schatzlein, C.E.
submitted to the EMBL Data Library, June 1995
A;Reference number: S57565
A;Accession: S57565
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-281 <SCH>
A;Cross-references: UNIPARC:UPI000000D91A; EMBL:X89102; NID:G887455; PID:G887456
R;Alderson, M.R.; Tough, T.W.; Davis-Smith, T.; Braddy, S.; Falk, B.; Schooley, K.A.; Go
J. Exp. Med. 181, 71-77, 1995
A;Title: Fas ligand mediates activation-induced cell death in human T lymphocytes.
A;Reference number: I38554; MUID:95105731; PMID:7528780
A;Accession: I38554
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-281 <RE2>
A;Cross-references: UNIPARC:UPI000000D91A; EMBL:U08137; NID:G624627; PIDN:AAC50071.1; PI
C;Genetics:
A;Gene: FasL
A;Introns: 151/1; 116/3
C;Keywords: glycoprotein; transmembrane protein
F;80-102/Domain: transmembrane #status predicted <TWM>
F;76,184,250,260/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 82.1%; Score 55; DB 2; Length 281;
Best Local Similarity 80.0%; Pred. No. 9.4;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPYPPPPPIP 10
||| |||||:
Db 46 PPPPPPPPLP 55

RESULT 15
T48514
hypothetical protein F15N18.140 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 09-Jul-2004
C;Accession: T48514
R;Bayan, M.; Hilbert, H.; Braun, M.; Holzer, E.; Brandt, A.; Duesterhoeft, A.; Bancroft,
submitted to the Protein Sequence Database, April 2000
A;Reference number: Z24490
A;Accession: T48514
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-314 <BEV>
A;Cross-references: UNIPROT:Q9LYD7; UNIPARC:UPI000000BE72; EMBL:AL163815
A;Experimental source: cultivar Columbia; BAC clone F15N18
C;Genetics:
A;Map position: 5
A;Note: F15N18.140

Query Match 82.1%; Score 55; DB 2; Length 314;
Best Local Similarity 80.0%; Pred. No. 10;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPYPPPPPIP 10
||| |||||:
Db 81 PPPPPPPPPP 90

RESULT 16
S02193
cellular tumor antigen p53 - chicken
N;Alternate names: nuclear oncoprotein p53
C;Species: Gallus gallus (chicken)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: S02193
R;Soussi, T.; Begue, A.; Kress, M.; Stehelin, D.; May, P.
Nucleic Acids Res. 16, 11383, 1988
A;Title: Nucleotide sequence of a cDNA encoding the chicken p53 nuclear oncoprotein.

A;Reference number: S02193; MUID:89083584; PMID:3060861
A;Accession: S02193
A;Molecule type: mRNA
A;Residues: 1-367 <SOU>
A;Cross-references: UNIPROT:P10360; UNIPARC:UPI000013103A; EMBL:X13057; NID:G63740; PIDN:
C;Superfamily: cellular tumor antigen p53
C;Keywords: apoptosis; cell division control; DNA binding; homotetramer; nucleus; phospho
F;161,164,224,228/Binding site: zinc (Cys, His, Cys, Cys) #status predicted
F;366/Binding site: phosphoryl-RNA (Ser) (covalent) #status predicted

Query Match 82.1%; Score 55; DB 1; Length 367;
Best Local Similarity 80.0%; Pred. No. 12;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPYPPPPPIP 10
||| |||||:
Db 50 PPPPPPPPLP 59

RESULT 17
T07910
hydroxyproline-rich glycoprotein GAS29 precursor - Chlamydomonas reinhardtii
C;Species: Chlamydomonas reinhardtii
C;Date: 21-May-1999 #sequence_revision 21-May-1999 #text_change 09-Jul-2004
C;Accession: T07910
R;Rodriguez-Martinez, H.; Haring, M.A.; Von Gromof, E.; Beck, C.F.
submitted to the EMBL Data Library, July 1997
A;Reference number: Z16207
A;Accession: T07910
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-433 <ROD>
A;Cross-references: UNIPROT:O22459; UNIPARC:UPI000000A1D78; EMBL:AF015884; NID:G2384729;
A;Experimental source: gametes
C;Genetics:
A;Gene: GAS29
F;1-31/Domain: signal sequence #status predicted <SIG>
F;32-433/Product: hydroxyproline-rich glycoprotein GAS29 #status predicted <MAT>

Query Match 82.1%; Score 55; DB 2; Length 433;
Best Local Similarity 80.0%; Pred. No. 15;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPYPPPPPIP 10
||| |||||:
Db 63 PPPPPPPPLP 72

RESULT 18
T34018
hypothetical protein Y4C6B.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C;Accession: T34018
R;Layman, D.; Graves, T.; Yoakum, M.
submitted to the EMBL Data Library, February 1999
A;Description: The sequence of C. elegans cosmid Y4C6B.
A;Reference number: Z21460
A;Accession: T34018
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-442 <LAY>
A;Cross-references: UNIPROT:Q9UB01; UNIPARC:UPI00000074EDD; EMBL:AF125971; PIDN:AAD14764.
A;Experimental source: strain Bristol N2; clone Y4C6B
C;Genetics:
A;Gene: CESP:Y4C6B.1
A;Map position: 4
A;Introns: 43/3; 78/3; 268/3; 387/2
C;Superfamily: Caenorhabditis elegans hypothetical protein Y4C6B.1

Query Match 82.1%; Score 55; DB 2; Length 442;
Best Local Similarity 80.0%; Pred. No. 15;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPYPPPPPIP 10
||| |||||:
Db 315 PPPPPPPPLP 324

RESULT 19
B85187
C:Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C/Accession: B85187
R:anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring Nature 402, 769-777, 1999
A/Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A/Reference number: A85001; MUID:20083488; PMID:10617198
A/Accession: B85187
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-473 <STO>
A/Cross-references: UNIPROT:Q98UK6; UNIPARC:UPI000000AC27B; GB:NC_001269; NID:g7268429; F:
C/Genetics:
A/Gene: AT4g16790
A/Map position: 4

Query Match 82.1%; Score 55; DB 2; Length 473;
Best Local Similarity 80.0%; Pred. No. 16;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPYPPPPPIP 10
||| |||||:
Db 284 PPPPPPPPLP 293

RESULT 20
E97813
WASP, N-WASP, MENA proteins homolog [imported] - Rickettsia conorii (strain Malish 7)
C/Species: Rickettsia conorii
C/Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C/Accession: E97813
R:Ogate, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.; R Science 293, 2093-2098, 2001
A/Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.
A/Reference number: A97700; MUID:21442074; PMID:11557893
A/Accession: E97813
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-520 <KUR>
A/Cross-references: UNIPROT:Q92H62; UNIPARC:UPI000000D4EF6; GB:AE006914; PIDN:AAL03447.1;
C/Genetics:
A/Gene: RC0909

Query Match 82.1%; Score 55; DB 2; Length 520;
Best Local Similarity 80.0%; Pred. No. 17;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPYPPPPPIP 10
||| |||||:
Db 352 PPPPPPPPLP 361

RESULT 21
I38607
p53-binding protein 2 - human (fragment)
C/Species: Homo sapiens (man)
C/Date: 09-Mar-1996 #sequence_revision 09-Mar-1996 #text_change 31-Dec-2004
C/Accession: I38607
R:Iwabuchi, K.; Bartel, P.L.; Li, B.; Marracino, R.; Fields, S. Proc. Natl. Acad. Sci. U.S.A. 91, 6098-6102, 1994
A/Title: Two cellular proteins that bind to wild-type but not mutant p53.
A/Reference number: I38604; MUID:94286584; PMID:18016121
A/Accession: I38607
A/Status: preliminary; nucleic acid sequence not shown

A/Molecule type: mRNA
A/Residues: 1-529 <RES>
A/Cross-references: UNIPROT:Q13625; UNIPARC:UPI000016A0A8; EMBL:U09582; NID:g493079; PIDN:
F:392-424/Domain: ankryrin repeat homology <AN08>
F:465-515/Domain: SH3 homology <SH3>

Query Match 82.1%; Score 55; DB 2; Length 529;
Best Local Similarity 88.9%; Pred. No. 18;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PPPYPPPPPIP 10
||| |||||:
Db 268 PPPYPPPPYP 276

RESULT 22
D40228
neurexin II-beta precursor - rat
C/Species: Rattus norvegicus (Norway rat)
C/Date: 21-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 09-Jul-2004
C/Accession: D40228; S27888
R:Ushkaryov, Y.A.; Petrenko, A.G.; Geppert, M.; Suedhof, T.C. Science 257, 50-56, 1992
A/Title: Neurexins: synaptic cell surface proteins related to the alpha-latrotoxin receptor
A/Reference number: A40228; MUID:92320296; PMID:1621094
A/Accession: D40228
A/Status: preliminary; nucleic acid sequence not shown
A/Molecule type: mRNA
A/Residues: 1-662 <USH>
A/Cross-references: UNIPROT:Q63376; UNIPARC:UPI0000130AA5; GB:M96377; NID:g205717; PIDN:
A/Note: authors translated GAC for residue 411 as Thr and ACC for residue 412 as Asp
C/Superfamily: neurexin; EGF homology
C/Keywords: alternative splicing; transmembrane protein
F:1-47/Domain: signal sequence #status predicted <SIG>
F:48-662/Product: neurexin II-beta #status predicted <MAT>

Query Match 82.1%; Score 55; DB 2; Length 662;
Best Local Similarity 80.0%; Pred. No. 22;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPYPPPPPIP 10
||| |||||:
Db 23 PPPPPPPPLP 32

RESULT 23
JC5929
serine/arginine-rich protein-specific kinase (BC 2.-.-.-) 2 - mouse
C/Species: Mus musculus (house mouse)
C/Date: 10-Apr-1998 #sequence_revision 08-May-1998 #text_change 09-Jul-2004
C/Accession: JC5929
R:Kuroyanagi, N.; Onogi, H.; Wakabayashi, T.; Hagiwara, M. Biochem. Biophys. Res. Commun. 242, 357-364, 1998
A/Title: Novel SR-protein-specific kinase, SRPK2, disassembles nuclear speckles.
A/Reference number: JC5929; MUID:98113357; PMID:9446799
A/Accession: JC5929
A/Status: nucleic acid sequence not shown
A/Molecule type: mRNA
A/Residues: 1-681 <KUR>
A/Cross-references: UNIPROT:Q8VCD9; UNIPARC:UPI000017C748
A/Experimental source: brain
C/Comment: This enzyme regulates the disassembly of the serine/arginine-rich proteins in a t
C/Keywords: transferase

Query Match 82.1%; Score 55; DB 2; Length 681;
Best Local Similarity 80.0%; Pred. No. 23;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPYPPPPPIP 10
||| |||||:
Db 27 PPPPPPPPLP 36

RESULT 24

T46476

hypothetical protein DKFPz434C0931.1 - human

C;Species: Homo sapiens (man)

C;Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Feb-2000

C;Accession: T46476

R;Blum, H.; Bauersachs, S.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, January 2000

A;Reference number: 223034

A;Accession: T46476

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-691 <AAA>

A;Cross-references: UNIPARC:UPI000015EC7C; EMBL:AL137718

A;Experimental source: adult testis; clone DKFPz434C0931

C;Genetics:

A;Note: DKFPz434C0931.1

Query Match 82.1%; Score 55; DB 2; Length 691;
Best Local Similarity 80.0%; Pred. No. 23;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPYPPPPIP 10
||| |||||
Db 328 PPPPPPPPLP 337

RESULT 25

D96563

probable bZIP protein, 48652-45869 [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004

C;Accession: D96563

R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Huizar, L.

Nature 408, 816-820, 2000

A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A;Reference number: A86141; MUID:21016719; PMID:11130712

A;Accession: D96563

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-798 <STO>

A;Cross-references: UNIPROT:Q9C824; UNIPARC:UPI00000A2A93; GB:AE005173; NID:gl0645444; F

C;Genetics:

A;Gene: F19K6.7

A;Map position: 1

Query Match 82.1%; Score 55; DB 2; Length 798;
Best Local Similarity 80.0%; Pred. No. 27;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPYPPPPIP 10
||| |||||
Db 116 PPPPPPPPLP 125

RESULT 26

E96636

hypothetical protein T7P1.21 [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004

C;Accession: E96636

R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Huizar, L.

Nature 408, 816-820, 2000

A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A;Reference number: A86141; MUID:21016719; PMID:11130712

A;Accession: E96636

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-907 <STO>

A;Cross-references: UNIPROT:Q9C946; UNIPARC:UPI00000A9620; GB:AE005173; NID:g6751696; P1:

C;Genetics:

A;Gene: T7P1.21

A;Map position: 1

Query Match 82.1%; Score 55; DB 2; Length 907;
Best Local Similarity 80.0%; Pred. No. 31;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPYPPPPIP 10
||| |||||
Db 508 PPPPPPPPLP 517

RESULT 27

S19595

chloride channel protein - rat

C;Species: Rattus norvegicus (Norway rat)

C;Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004

C;Accession: S19595

R;Steinmeyer, K.; Ortland, C.; Jentsch, T.J.

Nature 354, 301-304, 1991

A;Title: Primary structure and functional expression of a developmentally regulated skel:

A;Reference number: S19595; MUID:92065954; PMID:1659664

A;Accession: S19595

A;Molecule type: mRNA

A;Residues: 1-994 <STF>

A;Cross-references: UNIPROT:P35524; UNIPARC:UPI0000127A8F; GB:X62894; NID:g57744; PIDN: C

C;Keywords: transmembrane protein

F;830-877/Domain: CBS homology <CBS>

Query Match 82.1%; Score 55; DB 2; Length 994;
Best Local Similarity 80.0%; Pred. No. 34;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPYPPPPIP 10
||| |||||
Db 728 PPPPPPPPLP 737

RESULT 28

T13286

cappuccino gene protein - fruit fly (Drosophila melanogaster)

C;Species: Drosophila melanogaster

C;Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 09-Jul-2004

C;Accession: T13286

R;Emmons, S.; Phan, H.; Calley, J.; Chen, W.; James, B.; Manseau, L.

Genes Dev. 9, 2482-2494, 1995

A;Title: Cappuccino, a Drosophila maternal effect gene required for polarity of the egg

A;Reference number: Z17651; MUID:96033799; PMID:7590229

A;Accession: T13286

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-1058 <EMM>

A;Cross-references: UNIPROT:Q24120; UNIPARC:UPI000016BA5C; EMBL:U34258; NID:gi061333; P1:

C;Genetics:

A;Gene: capu

A;Cross-references: FlyBase:FBgn0000256

Query Match 82.1%; Score 55; DB 2; Length 1058;
Best Local Similarity 80.0%; Pred. No. 36;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPVPPPPPIP 10
||| |||||:
Db 486 PPPPPPPPLP 495

RESULT 29

Tl3963
formin related protein, lymphocyte specific - mouse
C/Species: Mus musculus (house mouse)
C/Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C/Accession: Tl3963
R/Taniuchi, I.; Yamamoto, S.; Watanabe, T.
submitted to the EMBL Data Library, June 1997
A/Description: Lymphocyte specific formin related gene.
A/Reference number: Z17840
A/Accession: Tl3963
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-1064 <TAN>
A/Cross-references: UNIPROT:Q922V7; UNIPARC:UPI0000002989C; EMBL:AF006466; NID:G4101719;
C/Genetics:
A/Gene: Fr1

Query Match 82.1%; Score 55; DB 2; Length 1064;
Best Local Similarity 80.0%; Pred. No. 36;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPVPPPPPIP 10
||| |||||:
Db 514 PPPPPPPPLP 523

RESULT 30

Tl7454
diaphanous-related formin - mouse
C/Species: Mus musculus (house mouse)
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C/Accession: Tl7454
R/Tominaga, T.; Sahai, E.; Treisman, R.H.; Alberts, A.S.
submitted to the EMBL Data Library, September 1998
A/Reference number: Z18796
A/Accession: Tl7454
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-1171 <TOM>
A/Cross-references: UNIPROT:Q92207; UNIPARC:UPI00000029711; EMBL:AF094519; NID:G3845724;
C/Genetics:
A/Gene: Dia2

Query Match 82.1%; Score 55; DB 2; Length 1171;
Best Local Similarity 80.0%; Pred. No. 40;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPVPPPPPIP 10
||| |||||:
Db 571 PPPPPPPPLP 580

RESULT 31

G86441
unknown protein [imported] - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C/Accession: G86441
R/Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A/Reference number: A86141; MUID:21016719; PMID:11130712
A/Accession: G86441
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-1201 <STO>
A/Cross-references: UNIPROT:Q9C6S1; UNIPARC:UPI0000009C9CB; GB:AB005172; NID:G11136725; P
C/Genetics:
A/Map position: 1

Query Match 82.1%; Score 55; DB 2; Length 1201;
Best Local Similarity 80.0%; Pred. No. 41;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPVPPPPPIP 10
||| |||||:
Db 574 PPPPPPPPLP 583

RESULT 32

S24407
formin isoform IV - mouse
C/Species: Mus musculus (house mouse)
C/Date: 19-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
C/Accession: S24407
R/Jackson-Grusby, L.; Kuo, A.; Leder, P.
Genes Dev. 6, 29-37, 1992
A/Title: A variant limb deformity transcript expressed in the embryonic mouse limb define
A/Reference number: S24407; MUID:92112033; PMID:1339380
A/Accession: S24407
A/Molecule type: mRNA
A/Residues: 1-1206 <JAC>
A/Cross-references: UNIPROT:Q05859; UNIPARC:UPI0000027920; EMBL:X62379; NID:G51552; PIDN

Query Match 82.1%; Score 55; DB 2; Length 1206;
Best Local Similarity 80.0%; Pred. No. 41;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPVPPPPPIP 10
||| |||||:
Db 705 PPPPPPPPLP 714

RESULT 33

T31065
diaphanous protein homolog p140mDia - mouse
C/Species: Mus musculus (house mouse)
C/Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 09-Jul-2004
C/Accession: T31065
R/Watanabe, N.; Madaule, P.; Reid, T.; Ishizaki, T.; Watanabe, G.; Kakizuka, A.; Saito, Y.
EMBO J. 16, 3044-3056, 1997
A/Title: p140mDia, a mammalian homolog of Drosophila diaphanous, is a target protein for
A/Reference number: Z20961; MUID:97357293; PMID:9214622
A/Accession: T31065
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-1255 <WAT>
A/Cross-references: UNIPROT:O08808; UNIPARC:UPI000002787C; EMBL:U96963; NID:G2114472; PII
A/Note: binds to GTP-bound form of Rho and binds to profilin

Query Match 82.1%; Score 55; DB 2; Length 1255;
Best Local Similarity 80.0%; Pred. No. 43;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPVPPPPPIP 10
||| |||||:
Db 586 PPPPPPPPLP 595

RESULT 34

S48375
hypothetical protein YIL159w - yeast (Saccharomyces cerevisiae)
C/Species: Saccharomyces cerevisiae

C;Date: 02-Dec-1994 #sequence_revision 02-Dec-1994 #text_change 09-Jul-2004

C;Accession: S48375

R;Churche, C.

submitted to the EMBL Data Library, September 1994

A;Reference number: S48310

A;Accession: S48375

A;Molecule type: DNA

A;Residues: 1-1375 <CHU>

A;Cross-references: UNIPROT:P40450; UNIPARC:UPI0000126A45; GB:Z47047; EMBL:Z38059; NID:9

C;Genetics:

A;Gene: SGD:BNR1

A;Cross-references: SGD:S0001421; MIPS:YIL159w

A;Map position: 9L

Query Match 82.1%; Score 55; DB 2; Length 1375;

Best Local Similarity 80.0%; Pred. No. 47;

Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPYPPPPIP 10

Db 772 PPPPPPPPLP 781

RESULT 35

S11515

formin - mouse

C;Species: Mus musculus (house mouse)

C;Date: 22-Jan-1994 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004

C;Accession: S11515

R;Woychik, R.P.; Maas, R.L.; Zeller, R.; Vogt, T.F.; Leder, P.

Nature 346, 850-853, 1990

A;Title: 'Formins': proteins deduced from the alternative transcripts of the limb deform

A;Reference number: S11515; MUID:90363291; PMID:2392150

A;Accession: S11515

A;Molecule type: mRNA

A;Residues: 1-1468 <WOY>

A;Cross-references: UNIPROT:Q05860; UNIPARC:UPI0000027927; EMBL:X53599; NID:G52877; PIDN

Query Match 82.1%; Score 55; DB 2; Length 1468;

Best Local Similarity 80.0%; Pred. No. 50;

Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPYPPPPIP 10

Db 931 PPPPPPPPLP 940

RESULT 36

F97713

190K antigen precursor [imported] - Rickettsia conorii (strain Malish 7)

C;Species: Rickettsia conorii

C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004

C;Accession: F97713

R;Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.; R

Science 293, 2093-2098, 2001

A;Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.

A;Reference number: A97700; MUID:21442074; PMID:11557893

A;Accession: F97713

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-1795 <KUR>

A;Cross-references: UNIPROT:Q92JF7; UNIPARC:UPI000000CBAD; GB:AE006914; PIDN:AAL02648.1;

C;Genetics:

A;Gene: sca2

Query Match 82.1%; Score 55; DB 2; Length 1795;

Best Local Similarity 80.0%; Pred. No. 61;

Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPYPPPPIP 10

Db 1078 PPPPPPPPLP 1087

RESULT 37

A41948

alpha-fetoprotein enhancer-binding protein - human

N;Alternate names: ATBF1 protein

C;Species: Homo sapiens (man)

C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 15-Oct-1999

A;Accession: A41948

R;Morinaga, T.; Yasuda, H.; Hashimoto, T.; Higashio, K.; Tamaoki, T.

Mol. Cell. Biol. 11, 6041-6049, 1991

A;Title: A human alpha-fetoprotein enhancer-binding protein, ATBF1, contains four homeod

A;Reference number: A41948; MUID:92049333; PMID:1719379

A;Accession: A41948

A;Molecule type: mRNA

A;Residues: 1-2783 <MOR>

A;Cross-references: UNIPARC:UPI0000156381; GB:D10250; GB:D90395; NID:G219429; PIDN:BA001

A;Note: sequence extracted from NCBI backbone (NCBIN:66271, NCBI:66276)

C;Genetics:

A;Gene: GDB:ATBF1

A;Cross-references: GDB:392090; OMIM:104155

A;Map position: 16q22.3-16q23.1

C;Superfamily: alpha-fetoprotein enhancer-binding protein; homeobox homology

C;Keywords: DNA binding; homeobox; nucleus; transcription regulation; zinc finger

F;72-94/Region: zinc finger CCHH motif

F;128-150/Region: zinc finger CCHH motif

F;176-198/Region: zinc finger CCHH motif

F;311-332/Region: zinc finger CCHH motif

F;340-361/Region: zinc finger CCHH motif

F;448-471/Region: zinc finger CCHH motif

F;489-509/Region: zinc finger CCHH motif

F;517-538/Region: zinc finger CCHH motif

F;633-655/Region: zinc finger CCHH motif

F;684-706/Region: zinc finger CCHH motif

F;719-773/Region: serine/threonine-rich

F;809-958/Region: glutamine-rich

F;1071-1092/Region: zinc finger CCHH motif

F;1117-1211/Region: proline-rich

F;1232-1288/Domain: homeobox homology <HOX1>

F;1329-1385/Domain: homeobox homology <HOX2>

F;1416-1437/Region: zinc finger CCHH motif

F;1618-1638/Region: zinc finger CCHH motif

F;1728-1784/Domain: homeobox homology <HOX3>

F;1799-1820/Region: zinc finger CCHH motif

F;2033-2089/Domain: homeobox homology <HOX4>

F;2112-2134/Region: zinc finger CCHH motif

F;2545-2566/Region: zinc finger CCHH motif

F;2585-2607/Region: glycine-rich

F;2611-2633/Region: zinc finger CCHH motif

F;2650-2737/Region: serine/threonine-rich

Query Match 82.1%; Score 55; DB 1; Length 2783;

Best Local Similarity 80.0%; Pred. No. 95;

Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPYPPPPIP 10

Db 1129 PPPPPPPPLP 1138

RESULT 38

T28957

hypothetical protein F45F2.11 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004

C;Accession: T28957

R;Davidson, S.; Wohldmann, P.

submitted to the EMBL Data Library, July 1996

A;Description: The sequence of C. elegans cosmid F45F2.

A;Reference number: Z20548

A;Accession: T28957

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-406 <DAV>

```
A;Cross-references: UNIPROT:Q29975; UNIPARC:UPI000007865A; EMBL:U64845; PIDN:AAC48032.1;
A;Experimental source: strain Bristol N2; clone F45P2
C;Genetics:
A;Gene: CESP:F45P2.11
A;Map position: 5
A;Introns: 1/2; 24/2; 143/3; 307/2; 338/2

Query Match      80.6%; Score 54; DB 2; Length 406;
Best Local Similarity 77.8%; Pred. No. 18;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 PPPPPPPP 10
   ||:||||:|
Db 109 PPPPPPLP 117

RESULT 39
T31502
hypothetical protein Y116A8C.32 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C;Accession: T31502
R;McMurray, A.
submitted to the EMBL Data Library, October 1999
A;Reference number: Z21041
A;Accession: T31502
A;Status: preliminary; translated from GB/EMBL/DBDJ
A;Molecule type: DNA
A;Residues: 1-710 <WIL>
A;Cross-references: UNIPROT:Q9U2U1; UNIPARC:UPI000017BC48; EMBL:AL117204; PIDN:CAB55136.
A;Experimental source: clone Y116A8C
C;Genetics:
A;Gene: CESP:Y116A8C.32
A;Introns: 53/1; 188/2; 238/3; 257/2; 308/2; 437/2; 600/1

Query Match      80.6%; Score 54; DB 2; Length 710;
Best Local Similarity 80.0%; Pred. No. 31;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPPYPPPP 10
   ||| |||||:|
Db 685 PPPPPPPMP 694

RESULT 40
T31170
diaphanous protein - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 09-Jul-2004
C;Accession: T31170
R;Castrillon, D.H.; Wasserman, S.A.; Castrillon, D.H.; Wasserman, S.A.
Development 120, 3367-3377, 1994
A;Title: Diaphanous is required for cytokinesis in Drosophila and shares domains of siml
A;Reference number: Z17626; MUID:95121197; PMID:7821209
A;Accession: T31170
A;Status: preliminary; translated from GB/EMBL/DBDJ
A;Molecule type: mRNA
A;Residues: 1-1091 <CAS>
A;Cross-references: UNIPROT:P48608; UNIPARC:UPI000016BA0C; EMBL:U11288; NID:9575926; PID
C;Genetics:
A;Gene: dia
A;Cross-references: FlyBase:FBgn0011202
A;Map position: 2L

Query Match      80.6%; Score 54; DB 2; Length 1091;
Best Local Similarity 80.0%; Pred. No. 48;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPPYPPPP 10
   ||| |||||:|
Db 539 PPPPPPPMP 548
```

```
RESULT 41
D48232
cysteine-rich extensin-like protein 4 precursor - common tobacco (fragment)
C;Species: Nicotiana tabacum (common tobacco)
C;Date: 26-May-1994 #sequence_revision 26-May-1994 #text_change 09-Jul-2004
C;Accession: D48232
R;Wu, H.; Zou, J.; May, B.; Gu, Q.; Cheung, A.Y.
Proc. Natl. Acad. Sci. U.S.A. 90, 6829-6833, 1993
A;Title: A tobacco gene family for flower cell wall proteins with a proline-rich domain
A;Reference number: A48232; MUID:93342083; PMID:8341705
A;Accession: D48232
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-160 <WUA>
A;Cross-references: UNIPROT:Q08197; UNIPARC:UPI0000177E6F; GB:L13442
A;Note: authors failed to translate the codon GCT for residue 11 as Ala, and GTA for res
A;Note: authors translated the codon TCT for residue 60 as Pro, and CCA for residue 117.
C;Genetics:
A;Gene: CELP-4
C;Superfamily: glutelin
C;Keywords: cell wall; extracellular matrix; fertilization; glycoprotein
F;1-17/Domain: signal sequence #status predicted <SIG>
F;18-159/Product: cysteine-rich extensin-like protein 4 #status experimental <MAT>

Query Match      79.1%; Score 53; DB 2; Length 160;
Best Local Similarity 80.0%; Pred. No. 8.9;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 PPPYPPPP 10
   ||| |||||
Db 74 PPPPPPPAP 83

RESULT 42
E48232
cysteine-rich extensin-like protein 5 precursor - common tobacco
C;Species: Nicotiana tabacum (common tobacco)
C;Date: 26-May-1994 #sequence_revision 26-May-1994 #text_change 09-Jul-2004
C;Accession: E48232
R;Wu, H.; Zou, J.; May, B.; Gu, Q.; Cheung, A.Y.
Proc. Natl. Acad. Sci. U.S.A. 90, 6829-6833, 1993
A;Title: A tobacco gene family for flower cell wall proteins with a proline-rich domain
A;Reference number: A48232; MUID:93342083; PMID:8341705
A;Accession: E48232
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-161 <WUA>
A;Cross-references: UNIPROT:Q08198; UNIPARC:UPI00000A99C3; GB:L13443; NID:9310930; PIDN:
C;Genetics:
A;Gene: CELP-5
C;Superfamily: glutelin
C;Keywords: cell wall; extracellular matrix; fertilization; glycoprotein
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-160/Product: cysteine-rich extensin-like protein 5 #status experimental <MAT>

Query Match      79.1%; Score 53; DB 2; Length 161;
Best Local Similarity 80.0%; Pred. No. 9;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 PPPYPPPP 10
   ||| |||||
Db 75 PPPPPPPAP 84

RESULT 43
C48232
cysteine-rich extensin-like protein 3 precursor - common tobacco
C;Species: Nicotiana tabacum (common tobacco)
C;Date: 26-May-1994 #sequence_revision 26-May-1994 #text_change 09-Jul-2004
C;Accession: C48232
R;Wu, H.; Zou, J.; May, B.; Gu, Q.; Cheung, A.Y.
Proc. Natl. Acad. Sci. U.S.A. 90, 6829-6833, 1993
A;Title: A tobacco gene family for flower cell wall proteins with a proline-rich domain
```

A;Reference number: A48232; MUID:93342083; PMID:8341705
A;Accession: C48232
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-165 <WUA>
A;Cross-references: UNIPROT:Q08196; UNIPARC:UPI00000AASD7; GB:L13441; NID:g310926; PIDN:
R;de S Goldman, M.H.; Pezzotti, M.; Seurinck, J.; Mariani, C.
A;Note: authors translated the codon GAT for residue 35 as Glu
C;Genetics:
A;Gene: CELP-3
C;Superfamily: glutelin

C;Keywords: cell wall; extracellular matrix; fertilization; glycoprotein
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-165/Product: cysteine-rich extensin-like protein 3 #status experimental <MAT>

Query Match 79.1%; Score 53; DB 2; Length 165;
Best Local Similarity 80.0%; Pred. No. 9.2; Mismatches 2; Indels 0; Gaps 0;
Matches 8; Conservative 0;

Qy 1 PPPYPPPPIP 10
||| |||||

Db 77 PPPSPPPPAP 86
||| |||||

RESULT 44

B48232

Cysteine-rich extensin-like protein 2 precursor - common tobacco

C;Species: Nicotiana tabacum (common tobacco)

C;Date: 26-May-1994 #sequence_revision 26-May-1994 #text_change 09-Jul-2004

C;Accession: B48232; PQ0474; S24616

R;Wu, H.; Zou, J.; May, B.; Gu, Q.; Cheung, A.Y.

Proc. Natl. Acad. Sci. U.S.A. 90, 6829-6833, 1993

A;Title: A tobacco gene family for flower cell wall proteins with a proline-rich domain

A;Reference number: A48232; MUID:93342083; PMID:8341705

A;Accession: B48232

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-196 <WUA>

A;Cross-references: UNIPROT:Q08195; UNIPARC:UPI00000A3B44; GB:L13440; NID:g310924; PIDN:
R;de S Goldman, M.H.; Pezzotti, M.; Seurinck, J.; Mariani, C.

Plant Cell 4, 1041-1051, 1992

A;Title: Developmental expression of tobacco pistil-specific genes encoding novel extens

A;Reference number: PQ0474; MUID:93005740; PMID:1392607

A;Accession: PQ0474

A;Molecule type: mRNA

A;Residues: 'MAG', 1-105 <GOL>

A;Cross-references: UNIPARC:UPI0000177B6E; EMBL:Z14014

A;Experimental source: stigma, style; strain Petit Havana SR1

C;Genetics:
A;Gene: CELP-2

C;Superfamily: glutelin

C;Keywords: cell wall; extracellular matrix; fertilization

F;1-19/Domain: signal sequence #status predicted <SIG>

F;20-196/Product: cysteine-rich extensin-like protein 2 #status experimental <MAT>

Query Match 79.1%; Score 53; DB 2; Length 196;
Best Local Similarity 80.0%; Pred. No. 11; Mismatches 2; Indels 0; Gaps 0;
Matches 8; Conservative 0;

Qy 1 PPPYPPPPIP 10
||| |||||

Db 107 PPPSPPPPAP 116
||| |||||

RESULT 45

A48232

Cysteine-rich extensin-like protein 1 precursor - common tobacco

C;Species: Nicotiana tabacum (common tobacco)

C;Date: 26-May-1994 #sequence_revision 26-May-1994 #text_change 09-Jul-2004

C;Accession: A48232; PQ0475; S24617

R;Wu, H.; Zou, J.; May, B.; Gu, Q.; Cheung, A.Y.

Proc. Natl. Acad. Sci. U.S.A. 90, 6829-6833, 1993

A;Title: A tobacco gene family for flower cell wall proteins with a proline-rich domain

A;Reference number: A48232; MUID:93342083; PMID:8341705

A;Accession: A48232

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-209 <WUA>

A;Cross-references: UNIPROT:Q08194; UNIPARC:UPI00000ABB3F; GB:L13439; NID:g310922; PIDN:
R;de S Goldman, M.H.; Pezzotti, M.; Seurinck, J.; Mariani, C.

Plant Cell 4, 1041-1051, 1992

A;Title: Developmental expression of tobacco pistil-specific genes encoding novel extens

A;Reference number: PQ0474; MUID:93005740; PMID:1392607

A;Accession: PQ0475

A;Molecule type: mRNA

A;Residues: 39-209 <GOL>

A;Cross-references: UNIPARC:UPI000016DE8A; EMBL:Z14020; NID:g19918; PID:g19919

A;Experimental source: stigma, style; strain Petit Havana SR1

C;Superfamily: glutelin

C;Keywords: cell wall; extracellular matrix; fertilization; glycoprotein

F;1-19/Domain: signal sequence #status predicted <SIG>

F;20-209/Product: cysteine-rich extensin-like protein 1 #status experimental <MAT>

F;146/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 79.1%; Score 53; DB 2; Length 209;
Best Local Similarity 80.0%; Pred. No. 12; Mismatches 2; Indels 0; Gaps 0;
Matches 8; Conservative 0;

Qy 1 PPPYPPPPIP 10
||| |||||

Db 119 PPPSPPPPAP 128
||| |||||

RESULT 46

B86255

hypothetical protein [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004

C;Accession: B86255

R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

Chen, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;

anssen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A;Reference number: A86141; MUID:21016719; PMID:11130712

A;Accession: B86255

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-744 <STO>

A;Cross-references: UNIPROT:O65375; UNIPARC:UPI00000A65A5; GB:AE005172; NID:g3157926; PI

C;Genetics:
A;Map position: 1

Query Match 79.1%; Score 53; DB 2; Length 744;
Best Local Similarity 80.0%; Pred. No. 42; Mismatches 2; Indels 0; Gaps 0;
Matches 8; Conservative 0;

Qy 1 PPPYPPPPIP 10
||| |||||

Db 525 PPPSPPPPCP 534
||| |||||

RESULT 47

T32404

hypothetical protein R148.3 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004

C;Accession: T32404

R;Le, T.T.; Kemp, K.; Scheet, P.

submitted to the EMBL Data Library, September 1997

A;Description: The sequence of C. elegans cosmid R148.

A;Reference number: Z21161

A;Accession: T32404
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-1127 <LET>
A;Cross-references: UNIPROT:O17267; UNIPARC:UPI0000176698; EMBL:AF025467; PIDN:RAB71038.
A;Experimental source: strain Bristol N2; clone R148
C;Genetics:
A;Gene: CRSP:R148.3
A;Map position: 3
A;Introns: 11/3; 42/1; 87/3; 373/3; 544/2; 595/2; 1042/1
C;Superfamily: Caenorhabditis elegans hypothetical protein R148.3

Query Match 79.1%; Score 53; DB 2; Length 1127;
Best Local Similarity 80.0%; Pred. No. 64;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PPPYPPPPIP 10
||| |||||
Db 388 PPAPPAPPAP 397

RESULT 48
JC5765
inositol polyphosphate 5-phosphatase (EC 3.1.3.-) 2 - human
C;Species: Homo sapiens (man)
C;Date: 24-Jan-1998 #sequence_revision 13-Mar-1998 #text_change 09-Jul-2004
C;Accession: JC5765
R;Pesesse, X.; Daleu, S.; De Smedt, F.; Drayer, L.; Erneux, C.
Biochem. Biophys. Res. Commun. 239, 697-700, 1997
A;Title: Identification of a second SH2-domain-containing protein closely related to the
A;Reference number: JC5765; MUID:98042455; PMID:9367831
A;Accession: JC5765
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-1258 <PES>
A;Cross-references: UNIPROT:O15357; UNIPARC:UPI000003ABB7; GB:Y14385; NID:g2653423; PIDN:
C;Comment: This enzyme plays a role in metabolism of inositol phosphate and phosphatidyl
C;Keywords: phosphoric monoester hydrolase
F;21-117/Domain: SH2 homology <SH2>
F;427-729/Domain: catalytic #status predicted <CAT>
F;927-1171/Region: proline-rich

Query Match 79.1%; Score 53; DB 2; Length 1258;
Best Local Similarity 70.0%; Pred. No. 72;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPYPPPPIP 10
||| |||||
Db 1050 PPDPFPPPLP 1059

RESULT 49
T00381
KIAA0633 protein - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 09-Jul-2004
C;Accession: T00381
R;Ishikawa, K.; Nagase, T.; Suyama, M.; Miyajima, N.; Tanaka, A.; Kotani, H.; Nomura, N.
DNA Res. 5, 169-176, 1998
A;Title: Prediction of the coding sequences of unidentified human genes. X. The complete
A;Reference number: Z14142; MUID:98403880; PMID:9734811
A;Accession: T00381
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-1316 <ISH>
A;Cross-references: UNIPROT:O75128; UNIPARC:UPI0000070BAF; EMBL:AB014533; NID:g3327079;
C;Genetics:
A;Gene: KIAA0633

Query Match 79.1%; Score 53; DB 2; Length 1316;
Best Local Similarity 80.0%; Pred. No. 75;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PPPYPPPPIP 10
||| |||||
Db 394 PPQPppppSP 403

RESULT 50
PQ0452
extensin-like protein - Persian tobacco (strain S283) (fragment)
C;Species: Nicotiana glauca (Persian tobacco)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 05-Oct-2004
C;Accession: PQ0452
R;Chen, C.G.; Cornish, B.C.; Clarke, A.E.
Plant Cell 4, 1053-1062, 1992
A;Title: Specific expression of an extensin-like gene in the style of Nicotiana glauca.
A;Reference number: PQ0451; MUID:93005741; PMID:1392608
A;Accession: PQ0452
A;Molecule type: mRNA
A;Residues: 1-122 <CHS>
A;Cross-references: UNIPROT:Q41192; UNIPARC:UPI0000177CC3
A;Experimental source: style, strain S283

Query Match 77.6%; Score 52; DB 2; Length 122;
Best Local Similarity 80.0%; Pred. No. 8.8;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PPPYPPPPIP 10
||| |||||
Db 54 PPSPppppSP 63

Search completed: April 6, 2006, 09:39:52
Job time : 32.7895 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 6, 2006, 09:23:24 ; Search time 145.895 Seconds
(without alignments)
58.030 Million cell updates/sec

Title: US-10-632-388-312

Perfect score: 67

Sequence: 1 PPPYPPPIPX 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Database : UniProt 05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	59	88.1	480	2	Q740X0 MYCPA
2	58	86.6	347	2	Q62127_ORYSA
3	58	86.6	467	2	Q62TK4 HUMAN
4	58	86.6	536	2	Q6ZVU3 HUMAN
5	58	86.6	539	2	Q5ZHV9 CHICK
6	58	86.6	559	1	3BP2 MOUSE
7	58	86.6	559	2	Q5U3L0 MOUSE
8	58	86.6	559	2	Q91252 MOUSE
9	58	86.6	559	2	Q8C220 MOUSE
10	58	86.6	561	1	3BP2 HUMAN
11	58	86.6	1166	2	Q7PVW3 ANOGA
12	58	86.6	1428	2	Q6Z128_ORYSA
13	57	85.1	172	2	Q00487_9CNID
14	57	85.1	222	2	Q6AVF5_ORYSA
15	57	85.1	231	2	Q5AX03_EMENI
16	57	85.1	282	1	ATF5 HUMAN
17	57	85.1	434	2	Q81W12 HUMAN
18	57	85.1	442	2	Q5VUC0 HUMAN
19	57	85.1	521	2	Q61WU8_CABBR
20	57	85.1	590	2	Q5SIV4_THET8
21	57	85.1	802	1	ENAH MOUSE
22	57	85.1	820	2	Q48682_ARATH
23	57	85.1	955	1	FRU DROME
24	57	85.1	1134	2	Q96JH1 HUMAN
25	57	85.1	1146	2	Q6LP66_CHLRE
26	57	85.1	1273	2	Q4WGB0_ASPFU
27	57	85.1	1324	2	Q8MMQ2_DICDI
28	57	85.1	1395	2	Q7SC01_NEUCR
29	57	85.1	1400	2	Q9VDD2_DROME
30	57	85.1	1918	2	Q752A6_ASHGO
31	56	83.6	135	2	Q5JM22_ORYSA

32	56	83.6	170	2	Q7YTN7_CABEL
33	56	83.6	173	2	Q61FU6_CABBR
34	56	83.6	218	2	Q52GY4_CABBR
35	56	83.6	233	2	Q64M77_ORYSA
36	56	83.6	247	2	Q4P543_USTMA
37	56	83.6	353	2	Q5FVW5_XENTR
38	56	83.6	393	1	EVL RAT
39	56	83.6	393	2	Q6PB99_MOUSE
40	56	83.6	405	2	Q501K3_MOUSE
41	56	83.6	409	2	Q9SBM1_VOLCA
42	56	83.6	414	1	EVL MOUSE
43	56	83.6	416	1	EVL HUMAN
44	56	83.6	418	2	Q7Z522_HUMAN
45	56	83.6	422	2	Q5R896_PONPY
46	56	83.6	437	2	Q7KK41_DROME
47	56	83.6	439	2	Q6YZ63_ORYSA
48	56	83.6	480	2	Q4R8F0_MACFA
49	56	83.6	489	2	Q73XK9_MYCPA
50	56	83.6	557	2	Q4R7B7_MACFA
51	56	83.6	563	2	Q9A9F1_CAUCR
52	56	83.6	582	2	Q5U118_DROME
53	56	83.6	582	2	Q18645_DROME
54	56	83.6	591	2	Q51L91_MAGGR
55	56	83.6	743	2	Q9BU60_HUMAN
56	56	83.6	744	1	1E63_HCMVA
57	56	83.6	809	2	Q76869_DROME
58	56	83.6	820	2	Q60XS1_CABBR
59	56	83.6	859	2	Q81RV3_DROME
60	56	83.6	887	2	Q9W4V1_DROME
61	56	83.6	954	2	Q4QF99_LEIMA
62	56	83.6	1009	2	Q8L685_VOLCA
63	56	83.6	1021	2	Q15451_HUMAN
64	56	83.6	1048	2	Q96FT1_HUMAN
65	56	83.6	1050	2	Q4NQ07_9DEL
66	56	83.6	1061	2	Q5EGN3_HUMAN
67	56	83.6	1130	2	Q81ZL8_HUMAN
68	56	83.6	1130	2	Q6NTE6_HUMAN
69	56	83.6	1284	2	Q15450_HUMAN
70	56	83.6	1493	2	Q9VEF7_DROME
71	56	83.6	1581	2	Q6PDJ2_MOUSE
72	56	83.6	1733	1	VNUA_PRVKA
73	56	83.6	1733	2	Q5PC68_9ALPH
74	56	83.6	1745	2	Q4QES4_LEIMA
75	56	83.6	1848	2	Q69ZN8_MOUSE
76	56	83.6	1953	1	BN11_YEAST
77	56	83.6	1958	2	Q69340_9ALPH
78	55	82.1	61	2	Q08549_MOUSE
79	55	82.1	71	2	Q54SX5_DICDI
80	55	82.1	73	2	Q7XIW2_ORYSA
81	55	82.1	89	2	Q7WXL8_ALCEU
82	55	82.1	103	2	Q84JC7_ORYSA
83	55	82.1	118	2	Q6YU92_ORYSA
84	55	82.1	121	2	Q6ZQN3_HUMAN
85	55	82.1	126	2	Q11299_MCV1
86	55	82.1	145	2	Q71KW7_RAT
87	55	82.1	147	2	Q9D2F0_MOUSE
88	55	82.1	163	2	Q9CU08_MOUSE
89	55	82.1	168	2	Q6ZPA5_HUMAN
90	55	82.1	169	2	Q35315_MOUSE
91	55	82.1	182	2	Q98324_MCV1
92	55	82.1	184	2	Q5VN30_ORYSA
93	55	82.1	187	2	Q6Z922_ORYSA
94	55	82.1	192	2	Q51SS5_MAGGR
95	55	82.1	199	2	Q84NQ7_ORYSA
96	55	82.1	203	2	Q7YYP5_CRYPV
97	55	82.1	209	2	Q8GRT1_ORYSA
98	55	82.1	210	2	Q9NXP9_HUMAN
99	55	82.1	211	2	Q6H016_FREDI
100	55	82.1	218	2	Q5HZL8_XENLA
101	55	82.1	218	2	Q9CUH8_MOUSE
102	55	82.1	229	2	Q6AVV7_ORYSA
103	55	82.1	239	2	Q9W2R5_DROME
104	55	82.1	242	2	Q559T7_DICDI

Q7YTN7	caenorhabdi
Q61FU6	caenorhabdi
Q52GY4	caenorhabdi
Q64M77	oryza sativ
Q4P543	utililago ma
Q5FVW5	xenopus tro
Q8R719	rattus norv
Q6PB99	mus musculus
Q501K3	mus musculus
Q9SBM1	volvox cart
P70429	mus musculus
Q9U108	homo sapien
Q7Z522	homo sapien
Q5R896	pongo pygma
Q7KK41	drosophila
Q6YZ63	oryza sativ
Q4R8F0	macaca fasc
Q73XK9	mycobacteri
Q4R7B7	macaca fasc
Q9A9F1	caulobacter
Q5U118	drosophila
Q18645	drosophila
Q51L91	magnaporthe
Q9BU60	homo sapien
P16749	human cytom
Q76869	drosophila
Q60XS1	caenorhabdi
Q81RV3	drosophila
Q9W4V1	drosophila
Q4QF99	leishmania
Q8L685	volvox cart
Q15451	homo sapien
Q96FT1	homo sapien
Q4NQ07	anaeromyxob
Q5EGN3	homo sapien
Q81ZL8	homo sapien
Q6NTE6	homo sapien
Q15450	homo sapien
Q9VEF7	drosophila
Q6PDJ2	mus musculus
P33485	pseudorabie
Q5PC68	suid herpes
Q4QES4	leishmania
Q69ZN8	mus musculus
P41832	saccharomyc
Q69340	suid herpes
Q08549	mus musculus
Q54SX5	dictyosteli
Q7XIW2	oryza sativ
Q7WXL8	alcaligenes
Q84JC7	oryza sativ
Q6YU92	oryza sativ
Q6ZQN3	homo sapien
Q11299	molluscum c
Q71KW7	rattus norv
Q9D2F0	mus musculus
Q9CU08	mus musculus
Q6ZPA5	homo sapien
Q35315	mus musculus
Q98324	molluscum c
Q5VN30	oryza sativ
Q6Z922	oryza sativ
Q51SS5	magnaporthe
Q84NQ7	oryza sativ
Q7YYP5	cryptospori
Q8GRT1	oryza sativ
Q9NXP9	homo sapien
Q6H016	fremyella d
Q5HZL8	xenopus lae
Q9CUH8	mus musculus
Q6AVV7	oryza sativ
Q9W2R5	drosophila
Q559T7	dictyosteli

105	55	82.1	243	2	Q75JU4_DICTDI	Q75JU4 dictyosteli	178	55	82.1	526	2	Q5XHx3_RAT	Q5XHx3 rattus norv
106	55	82.1	247	2	Q5W6v8_ORYSA	Q5w6v8 oryza sativ	179	55	82.1	532	2	Q9QY28_MOUSE	Q9QY28 mus musculus
107	55	82.1	262	1	YN03_CABEL	Q03607 caenorhabdi	180	55	82.1	533	2	Q80Zc2_MOUSE	Q80Zc2 mus musculus
108	55	82.1	268	2	Q9DGR3_CHICK	Q9dgn3 gallus gall	181	55	82.1	541	2	Q5D053_MOUSE	Q5D053 mus musculus
109	55	82.1	280	1	TNFL6_CERTO	Q9bdn1 c tumor nec	182	55	82.1	550	2	Q93263_CHICK	Q93263 gallus gall
110	55	82.1	280	1	TNFL6_MACFA	P63308 m tumor nec	183	55	82.1	558	2	Q6PIM5_HUMAN	Q6pim5 homo sapien
111	55	82.1	280	1	TNFL6_MACMU	P63307 m tumor nec	184	55	82.1	558	2	Q5BU88_XENTR	Q5bj88 xenopus tro
112	55	82.1	280	1	TNFL6_MACNE	P63306 m tumor nec	185	55	82.1	559	1	WASFL_HUMAN	Q92558 homo sapien
113	55	82.1	280	2	Q861W5_FELCA	Q861W5 felis silve	186	55	82.1	559	1	WASFL_MOUSE	Q85H6 mus musculus
114	55	82.1	281	1	TNFL6_HUMAN	P48023 h tumor nec	187	55	82.1	559	2	O5S2K7_HUMAN	Q5szk7 homo sapien
115	55	82.1	281	2	Q53Z2L_HUMAN	Q53z2l homo sapien	188	55	82.1	559	2	Q5NVG8_PONPY	Q5nv8 pongo pygma
116	55	82.1	283	2	Q4RDJ1_TETNG	Q4rdj1 tetraodon n	189	55	82.1	559	2	Q5BUJ7_RAT	Q5bju7 rattus norv
117	55	82.1	285	2	Q9LS54_ARATH	Q9ls54 arabidopsis	190	55	82.1	560	2	Q6GLN2_XENLA	Q6gln2 xenopus lae
118	55	82.1	285	2	Q6GY53_RAT	Q6gy53 rattus norv	191	55	82.1	568	2	Q96MU3_HUMAN	Q96mu3 homo sapien
119	55	82.1	285	2	Q6GY51_MOUSE	Q6gy51 mus musculus	192	55	82.1	579	2	Q4RLQ7_TETNG	Q4rlq7 tetraodon n
120	55	82.1	288	2	Q6XYB2_HUMAN	Q6xyb2 homo sapien	193	55	82.1	591	1	ENAH_HUMAN	Q8n87 homo sapien
121	55	82.1	289	2	Q8IXZ1_HUMAN	Q8ixz1 homo sapien	194	55	82.1	591	2	Q5QW5_HUMAN	Q5qz5 homo sapien
122	55	82.1	289	2	Q5BL00_HUMAN	Q5bl00 homo sapien	195	55	82.1	605	2	Q9Q5L3_9GAMA	Q9q5l3 cercopithec
123	55	82.1	290	1	HXDB_HUMAN	P13378 homo sapien	196	55	82.1	608	2	Q69S58_ORYSA	Q69s58 oryza sativ
124	55	82.1	296	2	Q5CW53_CRYPV	Q5cw53 cryptospori	197	55	82.1	631	2	Q8KQ9_MOUSE	Q8k2q9 mus musculus
125	55	82.1	306	2	O08551_MOUSE	O08551 mus musculus	198	55	82.1	634	2	Q8N9B5_HUMAN	Q8n9b5 homo sapien
126	55	82.1	307	2	Q4RS19_TETNG	Q4rs19 tetraodon n	199	55	82.1	649	2	Q68DG1_HUMAN	Q68dg1 homo sapien
127	55	82.1	314	2	Q9LYD7_ARATH	Q9lyd7 arabidopsis	200	55	82.1	662	1	NRX2B_RAT	Q83376 rattus norv
128	55	82.1	315	2	Q6ASR7_ORYSA	Q6asr7 oryza sativ	201	55	82.1	669	2	Q6MZFO_HUMAN	Q6mzf0 homo sapien
129	55	82.1	315	2	Q6GY52_RAT	Q6gy52 rattus norv	202	55	82.1	671	2	O95559_HUMAN	Q95559 homo sapien
130	55	82.1	321	2	Q6GY50_MOUSE	Q6gy50 mus musculus	203	55	82.1	673	2	Q4S708_TETNG	Q4s708 tetraodon n
131	55	82.1	321	2	O9PTD5_CHICK	O9ptd5 gallus gall	204	55	82.1	681	1	SRPK2_MOUSE	O54781 mus musculus
132	55	82.1	326	2	Q6UUS0_ORYSA	Q6uus0 oryza sativ	205	55	82.1	681	2	Q9VD05_DROME	Q9vd05 drosophila
133	55	82.1	328	1	MTFR1_MOUSE	Q9mb2 mus musculus	206	55	82.1	686	2	Q5DTW5_MOUSE	Q5dtw5 mus musculus
134	55	82.1	328	2	Q8CGA8_MOUSE	Q8cga8 mus musculus	207	55	82.1	687	2	Q4R9A2_MACFA	Q4r9a2 macaca fasc
135	55	82.1	333	2	Q99PM7_MOUSE	Q99pm7 mus musculus	208	55	82.1	688	1	SPRK2_HUMAN	Q78362 homo sapien
136	55	82.1	359	2	Q7FAP6_ORYSA	Q7fav6 oryza sativ	209	55	82.1	692	2	Q581R1_9TRYP	Q581r1 trypanosoma
137	55	82.1	359	2	Q7FAP4_ORYSA	Q7fap4 oryza sativ	210	55	82.1	692	2	Q4SUB2_TETNG	Q4sub2 tetraodon n
138	55	82.1	359	2	Q7XP48_ORYSA	Q7xp48 oryza sativ	211	55	82.1	693	2	Q4T7B3_TETNG	Q4t7b3 tetraodon n
139	55	82.1	360	2	Q96HE9_HUMAN	Q96he9 homo sapien	212	55	82.1	694	2	Q6F359_ORYSA	Q6f359 oryza sativ
140	55	82.1	360	2	Q9NUZ7_HUMAN	Q9nuz7 homo sapien	213	55	82.1	699	2	Q8NAV4_HUMAN	Q8nav4 homo sapien
141	55	82.1	367	1	P53_CHICK	P10360 gallus gall	214	55	82.1	723	2	O5T2S7_HUMAN	Q5t2s7 homo sapien
142	55	82.1	368	2	Q8BHE0_MOUSE	Q8bhe0 m riken cdn	215	55	82.1	724	2	Q80VK9_MOUSE	Q80vk9 mus musculus
143	55	82.1	420	2	Q9DEG1_CHICK	Q9deg1 gallus gall	216	55	82.1	735	2	Q4T2R8_TETNG	Q4t2r8 tetraodon n
144	55	82.1	427	2	Q6ETL2_ORYSA	Q6etl2 oryza sativ	217	55	82.1	737	2	Q5R541_PONPY	Q5re41 pongo pygma
145	55	82.1	433	2	Q22459_CHLRE	Q22459 chlamydomon	218	55	82.1	737	2	Q8H4U8_ORYSA	Q8h4u8 oryza sativ
146	55	82.1	435	2	Q9SM17_MAIZE	Q9sm17 zea mays (m	219	55	82.1	742	2	Q9FW12_ORYSA	Q9fw12 oryza sativ
147	55	82.1	435	2	Q80ZB7_RAT	Q80zb7 rattus norv	220	55	82.1	749	2	Q8NB24_HUMAN	Q8nb24 macaca fasc
148	55	82.1	442	2	Q9UB01_CABEL	Q9ub01 caenorhabdi	221	55	82.1	771	2	Q4R7B4_MACFA	Q4r7b4 macaca fasc
149	55	82.1	443	2	Q80ZC1_MOUSE	Q80zcl mus musculus	222	55	82.1	784	2	Q53S01_HUMAN	Q53sq1 homo sapien
150	55	82.1	444	2	Q80ZB6_RAT	Q80zb6 rattus norv	223	55	82.1	784	2	Q90YB5_CHICK	Q90y55 gallus gall
151	55	82.1	445	2	Q95KJ7_MACFA	Q95kj7 macaca fasc	224	55	82.1	795	2	Q5O4V9_HUMAN	Q5o4v9 homo sapien
152	55	82.1	446	2	Q9HCH4_HUMAN	Q9hch4 homo sapien	225	55	82.1	798	2	Q9C824_ARATH	Q9c824 arabidopsis
153	55	82.1	447	2	Q4SQR1_TETNG	Q4sqrl tetraodon n	226	55	82.1	798	2	Q571B6_MOUSE	Q571b6 mus musculus
154	55	82.1	451	2	Q62775_RAT	Q62775 rattus norv	227	55	82.1	824	2	Q8C414_MOUSE	Q8c414 mus musculus
155	55	82.1	452	2	Q80ZC0_MOUSE	Q80zc0 mus musculus	228	55	82.1	825	2	Q7S9H3_NEUCR	Q7s9h3 neurospora
156	55	82.1	456	2	Q8C4F6_MOUSE	Q8c4f6 mus musculus	229	55	82.1	830	2	Q6NRA4_XENLA	Q6nra4 xenopus lae
157	55	82.1	464	2	Q9FF14_ARATH	Q9ff14 arabidopsis	230	55	82.1	835	2	Q6PGC9_MOUSE	Q6pgc9 mus musculus
158	55	82.1	470	2	Q9LEM1_9BRAS	Q9lem1 capsella ru	231	55	82.1	836	2	Q4WU48_ASPFU	Q4wu48 aspergillus
159	55	82.1	471	2	Q9M3R6_ARATH	Q9m3r6 arabidopsis	232	55	82.1	837	2	Q63618_RAT	Q63618 rattus norv
160	55	82.1	473	2	Q8LCW1_ARATH	Q8lcw1 arabidopsis	233	55	82.1	844	2	Q6ZSV1_HUMAN	Q6zsv1 homo sapien
161	55	82.1	473	2	Q9SUK6_ARATH	Q9suk6 arabidopsis	234	55	82.1	849	2	Q6NUP0_HUMAN	Q6nu0 homo sapien
162	55	82.1	483	2	Q6ZK43_ORYSA	Q6zk43 oryza sativ	235	55	82.1	849	2	Q5XKF6_HUMAN	Q5xkf6 homo sapien
163	55	82.1	485	2	Q9Z0G8_RAT	Q9z0g8 rattus norv	236	55	82.1	853	1	D1AP3_HUMAN	Q9nv4 homo sapien
164	55	82.1	489	2	Q7YTX4_DROME	Q7ytx4 drosophila	237	55	82.1	853	2	Q4SUT0_TETNG	Q4sjt0 tetraodon n
165	55	82.1	494	2	Q96HR8_HUMAN	Q96hr8 homo sapien	238	55	82.1	854	2	Q5JXM8_HUMAN	Q5jxm8 homo sapien
166	55	82.1	494	2	Q9AKJ0_RICRI	Q9akj0 rickettsia	239	55	82.1	854	2	Q9H0A2_HUMAN	Q9h0a2 homo sapien
167	55	82.1	495	2	Q4RWK6_TETNG	Q4rwn6 tetraodon n	240	55	82.1	854	2	Q5T8W7_HUMAN	Q5t8w7 homo sapien
168	55	82.1	497	2	Q5RA03_PONPY	Q5ra03 pongo pygma	241	55	82.1	871	2	Q9ET47_MOUSE	Q9et47 mus musculus
169	55	82.1	501	2	Q4T4F3_TETNG	Q4t4f3 tetraodon n	242	55	82.1	884	2	Q9MA60_ARATH	Q9ma60 arabidopsis
170	55	82.1	502	2	Q7XRX7_ORYSA	Q7xrx7 oryza sativ	243	55	82.1	907	1	ANDR_CANFA	Q9tt90 canis famil
171	55	82.1	511	2	Q80ZB9_RAT	Q80zb9 rattus norv	244	55	82.1	922	2	Q9C946_ARATH	Q9c946 arabidopsis
172	55	82.1	513	2	Q9DEG2_CHICK	Q9deg2 gallus gall	245	55	82.1	927	2	Q7TT75_RAT	Q7tt75 rattus norv
173	55	82.1	517	2	Q8LN36_ORYSA	Q8ln36 oryza sativ	246	55	82.1	927	2	Q522P1_HUMAN	Q522p1 homo sapien
174	55	82.1	520	2	Q92H62_RICCN	Q92h62 rickettsia	247	55	82.1	929	2	Q8K331_MOUSE	Q8k331 mus musculus
175	55	82.1	520	2	Q80ZB8_RAT	Q80zb8 rattus norv	248	55	82.1	964	2	Q7TT74_RAT	Q7tt74 rattus norv
176	55	82.1	521	2	Q612S1_CAEBR	Q612s1 caenorhabdi	249	55	82.1	971	1	R3HDM_HUMAN	Q15032 homo sapien
177	55	82.1	524	2	Q80ZC3_MOUSE	Q80zc3 mus musculus	250	55	82.1	980	2	Q60PF6_CAEBR	Q60pf6 caenorhabdi

251	55	82.1	983	2	09QXN1_MOUSE	Q9qxm1 mus musculus
252	55	82.1	983	2	Q5BL16_MOUSE	Q5bl16 mus musculus
253	55	82.1	994	1	CLCN1_RAT	P35524 rattus norv
254	55	82.1	996	2	Q6CKU5_CRYHO	Q6ckj5 cryptospori
255	55	82.1	998	2	Q6GNN2_XENLA	Q6gmj2 xenopus lae
256	55	82.1	1009	2	Q8N3B0_HUMAN	Q8n3e0 homo sapien
257	55	82.1	1037	2	Q73808_FUGRU	Q73808 fugu rubrip
258	55	82.1	1051	2	Q9SE97_ATH	Q9se97 arabidopsis
259	55	82.1	1060	2	Q6NSN9_BRARE	Q6nsn9 brachydanio
260	55	82.1	1061	2	Q5R7K5_FONPY	Q5r7k5 pongo pygma
261	55	82.1	1087	2	Q59L24_CANAL	Q59l24 candida alb
262	55	82.1	1088	1	ASPP2_MOUSE	Q8cg79 mus musculus
263	55	82.1	1094	1	FMNL_MOUSE	Q9j126 mus musculus
264	55	82.1	1098	1	DIAP2_MOUSE	Q70566 mus musculus
265	55	82.1	1099	2	Q8IW32_HUMAN	Q8iw32 homo sapien
266	55	82.1	1102	2	Q6W4W7_MOUSE	Q6w4w7 mus musculus
267	55	82.1	1107	2	Q54Z23_DICDI	Q54z23 dictyosteli
268	55	82.1	1112	2	Q96PV5_HUMAN	Q6gpv5 homo sapien
269	55	82.1	1125	2	Q66IP5_XENLA	Q66ip5 xenopus lae
270	55	82.1	1128	1	ASPP2_HUMAN	Q13625 homo sapien
271	55	82.1	1130	2	Q9C0D6_HUMAN	Q9c0d6 homo sapien
272	55	82.1	1139	2	Q6NTK5_XENLA	Q6ntk5 xenopus lae
273	55	82.1	1141	2	Q4KLW6_XENLA	Q4klw6 xenopus lae
274	55	82.1	1145	2	Q4KLW4_XENLA	Q4klv4 xenopus lae
275	55	82.1	1147	2	Q5PR81_MOUSE	Q5pr81 mus musculus
276	55	82.1	1156	2	Q7XGQ4_ORYSA	Q7xgq4 oryza sativ
277	55	82.1	1156	2	Q8W2Q5_ORYSA	Q8w2q5 oryza sativ
278	55	82.1	1165	2	Q4T2J8_TETNG	Q4t2j8 tetraodon n
279	55	82.1	1167	2	Q640S7_XENTR	Q640s7 xenopus tro
280	55	82.1	1171	1	DIAP3_MOUSE	Q9z207 mus musculus
281	55	82.1	1175	2	Q4W6X2_COPCI	Q4w6x2 coprinus ci
282	55	82.1	1192	2	Q5DTM7_MOUSE	Q5dtm7 mus musculus
283	55	82.1	1201	2	Q9C6S1_ATH	Q9c6s1 arabidopsis
284	55	82.1	1204	2	Q4S986_TETNG	Q4s986 tetraodon n
285	55	82.1	1206	1	FMNLB_MOUSE	Q05859 mus musculus
286	55	82.1	1216	2	Q4SLQ6_TETNG	Q4slq6 tetraodon n
287	55	82.1	1220	2	Q6NS79_MOUSE	Q6ns79 mus musculus
288	55	82.1	1236	2	Q9C012_HUMAN	Q9c012 homo sapien
289	55	82.1	1248	1	DIAP1_MOUSE	Q06010 homo sapien
290	55	82.1	1250	2	Q70E73_HUMAN	Q70e73 homo sapien
291	55	82.1	1255	1	DIAP1_MOUSE	Q08808 mus musculus
292	55	82.1	1269	2	Q7XH56_ORYSA	Q7xh56 oryza sativ
293	55	82.1	1269	2	Q8W5K6_ORYSA	Q8w5k6 oryza sativ
294	55	82.1	1271	2	Q5KSL6_HUMAN	Q5ksl6 homo sapien
295	55	82.1	1272	2	Q6URQ4_HUMAN	Q6urc4 homo sapien
296	55	82.1	1278	2	Q4RQM1_TETNG	Q4rqm1 tetraodon n
297	55	82.1	1283	2	Q6CKG2_KLULA	Q6ckg2 kluyveromyc
298	55	82.1	1285	2	Q5DTQ4_MOUSE	Q5dtq4 mus musculus
299	55	82.1	1286	2	Q5BHU8_DROME	Q5bhu8 drosophila
300	55	82.1	1289	2	Q9FLQ7_ATH	Q9flq7 arabidopsis
301	55	82.1	1294	2	Q6FRQ4_CANGA	Q6frq4 candida gla
302	55	82.1	1299	2	Q59PF8_HUMAN	Q59pf8 homo sapien
303	55	82.1	1326	2	Q5XGV6_XENLA	Q5xgv6 xenopus lae
304	55	82.1	1347	2	Q4RW30_TETNG	Q4rw30 tetraodon n
305	55	82.1	1375	1	ENR1_YEAST	P40450 saccharomyc
306	55	82.1	1454	2	Q5SPR0_BRARE	Q5sep0 brachydanio
307	55	82.1	1466	2	Q56B22_RICAF	Q56b22 rickettsia
308	55	82.1	1468	1	FMNL1_MOUSE	Q05860 mus musculus
309	55	82.1	1527	1	ABHGB_RAT	Q9se67 rattus norv
310	55	82.1	1535	2	Q66JL8_MOUSE	Q66jl8 mus musculus
311	55	82.1	1552	2	Q68FM7_MOUSE	Q68fm7 mus musculus
312	55	82.1	1627	2	Q84ZL0_ORYSA	Q84zl0 oryza sativ
313	55	82.1	1635	2	Q7X693_ORYSA	Q7x693 oryza sativ
314	55	82.1	1641	2	Q5NCY0_MOUSE	Q5ncy0 mus musculus
315	55	82.1	1641	2	Q4VC26_MOUSE	Q4vc26 mus musculus
316	55	82.1	1652	2	Q9C0F0_HUMAN	Q9c0f0 homo sapien
317	55	82.1	1682	2	Q15054_HUMAN	Q15054 homo sapien
318	55	82.1	1707	2	Q6P9K2_MOUSE	Q6p9k2 mus musculus
319	55	82.1	1732	2	Q5AL52_CANAL	Q5al52 candida alb
320	55	82.1	1739	1	DOT1L_HUMAN	Q8tek3 homo sapien
321	55	82.1	1755	2	Q6BPB0_DEBHA	Q6bpe0 debaryomyce
322	55	82.1	1755	2	Q4IC42_GIBZE	Q4ic42 gibberella
323	55	82.1	1764	2	Q51V54_MAGGR	Q51vs4 magnaportha

324	55	82.1	1766	2	Q8NF45_HUMAN	Q8nf45 homo sapien
325	55	82.1	1795	2	Q92JF7_RICCN	Q92jf7 rickettsia
326	55	82.1	1835	2	Q6BNW7_DEBHA	Q6bnw7 debaryomyce
327	55	82.1	1842	2	Q6CUT9_KLULA	Q6cut9 kluyveromyc
328	55	82.1	1851	2	Q6C9I8_YARLI	Q6c9i8 yarrowia li
329	55	82.1	1883	2	Q8LMR2_ORYSA	Q8lmr2 oryza sativ
330	55	82.1	1884	2	Q4RWK6_TETNG	Q4rwk6 tetraodon n
331	55	82.1	1951	1	YLPW1_HUMAN	P49750 homo sapien
332	55	82.1	2068	2	Q67WF0_ORYSA	Q67wf0 oryza sativ
333	55	82.1	2147	2	Q61N21_CAEBR	Q61n21 caenorhabdi
334	55	82.1	2333	2	Q615U0_CAEBR	Q615u0 caenorhabdi
335	55	82.1	2333	2	Q86UP3_HUMAN	Q86up3 homo sapien
336	55	82.1	3567	1	Q86UP3_HUMAN	Q15911 homo sapien
337	55	82.1	3703	1	ATBFI_HUMAN	Q15911 homo sapien
338	55	82.1	3726	1	ATBFI_MOUSE	Q1329 mus musculus
339	55	82.1	4969	2	Q8CF9I_MOUSE	Q8cf9i mus musculus
340	55	82.1	5038	1	PCLO_MOUSE	Q9gyx7 mus musculus
341	55	82.1	5120	1	PCLO_CHICK	Q9qy36 gallus gall
342	55	82.1	5165	2	Q8CF92_MOUSE	Q8cf92 mus musculus
343	55	82.1	5183	1	PCLO_HUMAN	Q9y6v0 homo sapien
344	54.5	81.3	222	2	Q5TN00_ANOGA	Q5tn00 anopheles g
345	54.5	81.3	3070	2	Q4RLJ2_TETNG	Q4rlj2 tetraodon n
346	54	80.6	101	2	Q8BR19_MOUSE	Q8br19 mus musculus
347	54	80.6	129	2	Q5ZE72_ORYSA	Q5ze72 oryza sativ
348	54	80.6	192	2	Q7S944_NEUCR	Q7s944 neurospora
349	54	80.6	213	2	Q4KT55_9NUCL	Q4kt55 chrysodeixi
350	54	80.6	219	2	Q61UQ5_CAEBR	Q61uq5 caenorhabdi
351	54	80.6	352	2	Q6EF36_ORYSA	Q6ep36 oryza sativ
352	54	80.6	358	2	Q82239_ATH	Q82239 arabidopsis
353	54	80.6	389	2	Q6ET49_ORYSA	Q6et49 oryza sativ
354	54	80.6	401	2	Q9U3U1_CAEBR	Q9u3u1 caenorhabdi
355	54	80.6	406	2	Q22975_CAEBR	Q22975 caenorhabdi
356	54	80.6	414	2	Q624L9_CAEBR	Q624l9 caenorhabdi
357	54	80.6	417	2	Q4KT78_9NUCL	Q4kt78 chrysodeixi
358	54	80.6	462	2	Q9IBR4_9NUCL	Q9ibr4 spodoptera
359	54	80.6	485	1	PDIC7_HUMAN	Q8n8d1 homo sapien
360	54	80.6	515	2	Q7YTY2_DROME	Q7yty2 drosophila
361	54	80.6	515	2	Q9VCH2_DROME	Q9vch2 drosophila
362	54	80.6	522	2	Q60XF4_CAEBR	Q60xf4 caenorhabdi
363	54	80.6	664	2	Q62BE1_CAEBR	Q62be1 caenorhabdi
364	54	80.6	699	2	Q9U2U1_CAEBR	Q9u2u1 caenorhabdi
365	54	80.6	705	2	Q81VU2_HUMAN	Q81vu2 homo sapien
366	54	80.6	722	2	Q6ZVG1_HUMAN	Q6zvg1 homo sapien
367	54	80.6	732	2	Q8J1Y5_ASHGO	Q8j1y5 ashbya goos
368	54	80.6	732	2	Q74ZB4_ASHGO	Q74zb4 ashbya goos
369	54	80.6	788	2	Q69MT2_ORYSA	Q69mt2 oryza sativ
370	54	80.6	789	2	Q4SHS7_TETNG	Q4shs7 tetraodon n
371	54	80.6	1005	2	Q6ZVH7_HUMAN	Q6zvh7 homo sapien
372	54	80.6	1091	1	DIA_DROME	P48608 drosophila
373	54	80.6	1091	2	Q5BI26_DROME	Q5bi26 drosophila
374	54	80.6	1217	2	Q9ULL5_HUMAN	Q9ull5 homo sapien
375	54	80.6	1408	2	Q7Q976_ANOGA	Q7q976 anopheles g
376	54	80.6	1485	2	Q4I763_GIBZE	Q4i763 gibberella
377	54	80.6	1533	2	Q7SAT8_NEUCR	Q7sat8 neurospora
378	54	80.6	1612	2	Q7PV65_ANOGA	Q7pv65 anopheles g
379	54	80.6	1790	2	Q86ZG7_NEUCR	Q86zg7 neurospora
380	54	80.6	1817	2	Q7RWH7_NEUCR	Q7rwh7 neurospora
381	53	79.1	132	2	Q61H28_CAEBR	Q61h28 caenorhabdi
382	53	79.1	157	2	Q08197_TOBAC	Q08197 nicotiana t
383	53	79.1	161	2	Q08198_TOBAC	Q08198 nicotiana t
384	53	79.1	165	2	Q08196_TOBAC	Q08196 nicotiana t
385	53	79.1	171	2	Q7DLN2_TOBAC	Q7dlm2 nicotiana t
386	53	79.1	196	2	Q08195_TOBAC	Q08195 nicotiana t
387	53	79.1	209	2	Q08194_TOBAC	Q08194 nicotiana t
388	53	79.1	223	2	Q5VR17_ORYSA	Q5vr17 oryza sativ
389	53	79.1	235	2	Q02442_GLOBO	Q02442 globodera r
390	53	79.1	239	2	Q5VRU9_DROME	Q5vru9 drosophila
391	53	79.1	246	2	Q5VRU9_ORYSA	Q5vr4 oryza sativ
392	53	79.1	273	2	Q80YB9_MOUSE	Q80yb9 mus musculus
393	53	79.1	337	2	Q81VB5_HUMAN	Q81vb5 homo sapien
394	53	79.1	338	2	Q5PQ07_RAT	Q5pq07 rattus norv
395	53	79.1	347	2	Q585A8_9TRYP	Q585a8 trypanosoma
396	53	79.1	347	2	Q585B0_9TRYP	Q585b0 trypanosoma

397	53	79.1	348	2	Q8BQ89_MOUSE	Q8BQ89 mus musculus	470	52	77.6	243	2	Q67UD2_ORYSA	Q67UD2 oryza sativ
398	53	79.1	360	2	Q9XZM6_STRPU	Q9XZM6 strongyloce	471	52	77.6	247	2	Q5AY44_EMENI	Q5AY44 aspergillus
399	53	79.1	379	2	Q5O4V4_HUMAN	Q5O4V4 homo sapien	472	52	77.6	251	2	Q58WM7_9POAL	Q58WM7 elionurus t
400	53	79.1	400	2	Q9SQF7_BRAJU	Q9SQF7 braesica ju	473	52	77.6	251	2	Q8L3S3_ORYSA	Q8L3S3 oryza sativ
401	53	79.1	415	2	Q8S9B6_VOLCA	Q8S9B6 volvox carl	474	52	77.6	256	2	Q5YB71_MANSE	Q5YB71 manduca sex
402	53	79.1	479	2	Q08611_MOUSE	Q08611 mus musculus	475	52	77.6	256	2	Q58WL1_9POAL	Q58WL1 ischaemum a
403	53	79.1	491	2	Q7RXL8_NEUCR	Q7RXL8 neurospora	476	52	77.6	258	2	Q58WM3_9POAL	Q58WM3 coelotrachis
404	53	79.1	520	2	Q5VR46_ORYSA	Q5VR46 oryza sativ	477	52	77.6	258	2	Q58WM4_9POAL	Q58WM4 apluda muti
405	53	79.1	744	2	Q65375_ATH	Q65375 arabidopsis	478	52	77.6	258	2	Q58WM5_9POAL	Q58WM5 coelotrachis
406	53	79.1	899	2	Q8N304_HUMAN	Q8N304 homo sapien	479	52	77.6	260	2	Q58WK7_COILA	Q58WK7 coix lachry
407	53	79.1	1017	2	Q4SV87_TETNG	Q4SV87 tetraodon n	480	52	77.6	260	2	Q58WM2_9POAL	Q58WM2 coix aquatry
408	53	79.1	1068	2	Q6CCCO_YARLI	Q6CCCO yarrowia li	481	52	77.6	264	2	Q6VEQ4_HUMAN	Q6VEQ4 homo sapien
409	53	79.1	1087	2	Q7QIT1_ANOGA	Q7QIT1 anopheles g	482	52	77.6	264	2	Q6VEQ5_HUMAN	Q6VEQ5 homo sapien
410	53	79.1	1115	2	Q17267_CAENH	Q17267 caenorhabdi	483	52	77.6	264	2	Q6P447_HUMAN	Q6P447 homo sapien
411	53	79.1	1130	2	Q8ITY8_CAENH	Q8ITY8 caenorhabdi	484	52	77.6	264	2	Q4NYS5_9DELT	Q4NYS5 anaeromyxob
412	53	79.1	1146	2	Q86XA7_HUMAN	Q86XA7 homo sapien	485	52	77.6	268	1	LBD13_ATH	Q9AE61 arabidopsis
413	53	79.1	1149	2	Q13577_HUMAN	Q13577 homo sapien	486	52	77.6	270	2	Q7Z666_HUMAN	Q7Z666 homo sapien
414	53	79.1	1175	2	Q6A0A9_MOUSE	Q6A0A9 mus musculus	487	52	77.6	271	2	Q8NFH6_HUMAN	Q8NFH6 homo sapien
415	53	79.1	1183	2	Q9R1V2_RAT	Q9R1V2 rattus norv	488	52	77.6	275	2	Q7Z7K6_HUMAN	Q7Z7K6 homo sapien
416	53	79.1	1257	2	Q9JLL7_MOUSE	Q9JLL7 mus musculus	489	52	77.6	297	2	Q6ZT55_HUMAN	Q6ZT55 homo sapien
417	53	79.1	1257	2	Q9WVR3_RAT	Q9WVR3 rattus norv	490	52	77.6	297	2	Q67K16_SYMTH	Q67K16 symbiobacte
418	53	79.1	1257	2	Q6P549_MOUSE	Q6P549 mus musculus	491	52	77.6	299	2	Q7X838_ORYSA	Q7X838 oryza sativ
419	53	79.1	1258	2	Q15357_HUMAN	Q15357 homo sapien	492	52	77.6	300	2	Q4THH6_TETNG	Q4THH6 tetraodon n
420	53	79.1	1316	2	Q75128_HUMAN	Q75128 homo sapien	493	52	77.6	301	2	Q6L4D1_ORYSA	Q6L4D1 oryza sativ
421	53	79.1	1391	2	Q6K8Z4_ORYSA	Q6K8Z4 oryza sativ	494	52	77.6	302	2	Q5Z7H9_ORYSA	Q5Z7H9 oryza sativ
422	53	79.1	4027	2	Q512R0_CHLIN	Q512R0 chlamydomon	495	52	77.6	302	2	Q9CAC7_ATH	Q9CAC7 arabidopsis
423	52	77.6	39	2	Q9BR93_HUMAN	Q9BR93 homo sapien	496	52	77.6	305	2	Q9SH44_ATH	Q9SH44 arabidopsis
424	52	77.6	56	2	Q8MM17_ACRPL	Q8MM17 acropora pr	497	52	77.6	309	2	Q6IEL3_ORYSA	Q6IEL3 oryza sativ
425	52	77.6	56	2	Q8MM18_ACRPL	Q8MM18 acropora pa	498	52	77.6	309	2	Q7XNY3_ORYSA	Q7XNY3 oryza sativ
426	52	77.6	56	2	Q8MM19_ACRCE	Q8MM19 acropora ce	499	52	77.6	314	2	Q54817_MOUSE	Q54817 mus musculus
427	52	77.6	56	2	Q8MUZ3_ACRPL	Q8MUZ3 acropora pa	500	52	77.6	318	2	Q5PQY8_BRARE	Q5PQY8 brachydanio
428	52	77.6	56	2	Q8MUZ4_ACRPL	Q8MUZ4 acropora pa	501	52	77.6	323	1	LECT_SOLTU	Q98M0 solanum tub
429	52	77.6	56	2	Q8MUZ5_9CNID	Q8MUZ5 acropora na	502	52	77.6	329	2	Q9WX50_ACEXY	Q9WX50 acetobacter
430	52	77.6	74	2	Q59BP8_HUMAN	Q59BP8 homo sapien	503	52	77.6	333	2	Q7XR83_ORYSA	Q7XR83 oryza sativ
431	52	77.6	83	2	Q9GT19_HETGL	Q9GT19 heterodera	504	52	77.6	344	2	Q9H7P6_HUMAN	Q9H7P6 homo sapien
432	52	77.6	87	2	Q7WXH7_ALCEU	Q7WXH7 alcaligenes	505	52	77.6	351	2	Q7QD64_ANOGA	Q7QD64 anopheles g
433	52	77.6	90	2	Q4VMJ70_MANSE	Q4VMJ70 manduca sex	506	52	77.6	353	2	Q6TRW9_CHICK	Q6TRW9 gallus gall
434	52	77.6	115	2	Q5AL62_CANAL	Q5AL62 candida alb	507	52	77.6	356	2	Q6TRX0_CHICK	Q6TRX0 gallus gall
435	52	77.6	116	2	Q8C9A6_MOUSE	Q8C9A6 mus muscul	508	52	77.6	359	2	Q9XZT0_DROME	Q9XZT0 drosophila
436	52	77.6	137	2	Q4JMU1_9BACT	Q4JMU1 uncultured	509	52	77.6	361	2	Q6IMV8_ORYSA	Q6IMV8 oryza sativ
437	52	77.6	138	2	Q7QES1_ANOGA	Q7QES1 anopheles g	510	52	77.6	365	2	Q4KLX2_XENLA	Q4KLX2 xenopus lae
438	52	77.6	138	2	Q9S8K0_ATH	Q9S8K0 arabidopsis	511	52	77.6	367	2	Q59847_ASPOR	Q59847 aspergillus
439	52	77.6	144	2	Q23120_ATH	Q23120 arabidopsis	512	52	77.6	373	2	Q62CV6_BURMA	Q62CV6 burkholderi
440	52	77.6	144	2	Q8LAZ0_ATH	Q8LAZ0 arabidopsis	513	52	77.6	376	2	Q53N23_ORYSA	Q53N23 oryza sativ
441	52	77.6	145	2	Q8MLF7_DROME	Q8MLF7 drosophila	514	52	77.6	378	2	Q19961_CAENH	Q19961 caenorhabdi
442	52	77.6	147	2	Q8CBK2_MOUSE	Q8CBK2 mus muscul	515	52	77.6	378	2	Q6T9Z5_CHICK	Q6T9Z5 gallus gall
443	52	77.6	149	2	Q00484_9CNID	Q00484 hydra sp. m	516	52	77.6	382	2	Q88933_MOUSE	Q88933 mus muscul
444	52	77.6	149	2	Q35327_MOUSE	Q35327 mus muscul	517	52	77.6	383	2	Q60119_SCHPO	Q60119 schizosacch
445	52	77.6	151	1	YWIS_WHEAT	Y14723 triticum ae	518	52	77.6	387	2	Q9Z2U3_MOUSE	Q9Z2U3 mus muscul
446	52	77.6	151	2	Q41192_NICAL	Q41192 nicotiana a	519	52	77.6	391	2	Q8PXX6_BRUSU	Q8PXX6 bruceella su
447	52	77.6	153	1	PRIMA_HUMAN	Q86XR5 homo sapien	520	52	77.6	393	2	Q9N2Z3_CAENH	Q9N2Z3 caenorhabdi
448	52	77.6	153	1	PRIMA_MOUSE	Q810F0 mus muscul	521	52	77.6	393	2	Q91TI4_TUHV1	Q91TI4 tupaiid her
449	52	77.6	153	1	SORB_PIG	F28210 sus scrofa	522	52	77.6	396	2	Q5QIK4_MAIZE	Q5QIK4 zea mays (m
450	52	77.6	158	2	Q9SME6_PIG	Q9SME6 sus scrofa	523	52	77.6	396	2	Q5QIJ3_ZEAMP	Q5QIJ3 zea mays (s
451	52	77.6	158	2	Q9DVM0_9BACU	Q9DVM0 pluteella xy	524	52	77.6	396	2	Q86VE0_HUMAN	Q86VE0 homo sapien
452	52	77.6	162	2	Q8XSK0_RALSO	Q8XSK0 ralstonia s	525	52	77.6	399	2	Q9UPV1_HUMAN	Q9UPV1 homo sapien
453	52	77.6	170	2	Q8XTG7_RALSO	Q8XTG7 ralstonia s	526	52	77.6	402	2	Q710T9_POPE	Q710T9 populus del
454	52	77.6	176	2	Q16990_ACRDO	Q16990 acropora do	527	52	77.6	408	2	Q96RE2_HUMAN	Q96RE2 homo sapien
455	52	77.6	179	2	Q9D0I6_MOUSE	Q9D0I6 mus muscul	528	52	77.6	409	2	Q710T9_POPE	Q710T9 populus del
456	52	77.6	182	2	Q5M833_RAT	Q5M833 rattus norv	529	52	77.6	412	2	Q10610_9NUCL	Q10610 helicoverpa
457	52	77.6	186	2	Q00486_9CNID	Q00486 hydra sp. m	530	52	77.6	413	2	Q8V5X7_9NUCL	Q8V5X7 helicoverpa
458	52	77.6	187	2	Q67YU3_ATH	Q67YU3 arabidopsis	531	52	77.6	413	2	Q99H44_9NUCL	Q99H44 helicoverpa
459	52	77.6	192	2	Q4WN95_ASPFU	Q4WN95 aspergillus	532	52	77.6	418	2	Q5ZIC9_CHICK	Q5ZIC9 gallus gall
460	52	77.6	197	2	Q55BN3_DICDI	Q55BN3 dictyosteli	533	52	77.6	421	1	ACRO_HUMAN	P10323 homo sapien
461	52	77.6	206	2	Q825Z2_STRAW	Q825Z2 streptomyce	534	52	77.6	421	2	Q6ICK2_HUMAN	Q6ICK2 homo sapien
462	52	77.6	211	2	Q9M644_ATH	Q9M644 arabidopsis	535	52	77.6	428	2	Q23370_ATH	Q23370 arabidopsis
463	52	77.6	217	2	Q8IFN2_PLAF7	Q8IFN2 plasmodium	536	52	77.6	431	1	KNAT3_ATH	P48000 arabidopsis
464	52	77.6	217	2	Q8MS64_DROME	Q8MS64 drosophila	537	52	77.6	431	2	Q8LDC1_ATH	Q8LDC1 arabidopsis
465	52	77.6	221	2	Q8N6S7_HUMAN	Q8N6S7 homo sapien	538	52	77.6	442	2	Q8T9S2_DERVA	Q8T9S2 dermator
466	52	77.6	222	1	ACROL_HUMAN	P58840 homo sapien	539	52	77.6	450	2	Q7Z3N4_HUMAN	Q7Z3N4 homo sapien
467	52	77.6	241	2	Q58WN4_ZEALU	Q58WN4 zea luxuria	540	52	77.6	451	2	Q9SKL3_ATH	Q9SKL3 arabidopsis
468	52	77.6	241	2	Q8R2W2_MOUSE	Q8R2W2 mus muscul	541	52	77.6	452	2	Q4TB69_TETNG	Q4TB69 tetraodon n
469	52	77.6	242	2	Q58WN1_9POAL	Q58WN1 tripaecum a	542	52	77.6	453	2	Q86S65_CAENH	Q86S65 caenorhabdi

543	52	77.6	455	2	Q5AED9_CANAL	Q5aed9 candida alb	616	52	77.6	695	2	Q4R565_MACPA	Q4r565 macaca fasc
544	52	77.6	456	2	Q655D9_ORYSA	Q655q5 oryza sativ	617	52	77.6	698	2	Q7XR73_ORYSA	Q7xf73 oryza sativ
545	52	77.6	457	2	Q52KK4_RAT	Q52kk4 rattus norv	618	52	77.6	700	2	Q8S976_ORYSA	Q8s976 oryza sativ
546	52	77.6	464	2	Q41645_VOLCA	Q41645 volvox cart	619	52	77.6	707	2	Q653H7_ORYSA	Q653h7 oryza sativ
547	52	77.6	468	2	Q86TK9_HUMAN	Q86tk9 homo sapien	620	52	77.6	700	2	Q5RHD8_BRARE	Q5rhd8 brachydanio
548	52	77.6	471	2	Q503D7_BRARE	Q503d7 brachydanio	621	52	77.6	708	2	Q9LJG1_ARATH	Q9ljg1 arabidopsis
549	52	77.6	473	2	Q51RH2_MAGGR	Q51rh2 magnaporthe	622	52	77.6	708	2	Q8BR71_MOUSE	Q8br71 mus musculus
550	52	77.6	475	2	Q8VDD8_MOUSE	Q8vdd8 mus musculus	623	52	77.6	711	1	UBP51_HUMAN	Q8br71 mus musculus
551	52	77.6	475	2	Q6Y636_MOUSE	Q6y636 mus musculus	624	52	77.6	711	2	Q7S0F5_NEUCR	Q70ek9 homo sapien
552	52	77.6	477	2	Q85IM1_ORYSA	Q85im1 oryza sativ	625	52	77.6	712	2	Q52KFS_MOUSE	Q70ek9 mus musculus
553	52	77.6	482	2	Q6JV11_XENLA	Q6jv11 xenopus lae	626	52	77.6	724	1	KCNN3_PIG	Q52kfs mus musculus
554	52	77.6	487	1	EBN2_EBV	P12978 Epstein-Bar	627	52	77.6	727	2	Q4S6F4_TETNG	P58392 sus scrofa
555	52	77.6	487	2	Q777H1_9GAMA	Q777h1 human herpe	628	52	77.6	730	2	Q45021_CAEEL	Q45021 caenorhabdi
556	52	77.6	488	1	ETBD1_MOUSE	P58544 mus musculus	629	52	77.6	732	2	Q60XR9_CAEER	Q45021 caenorhabdi
557	52	77.6	488	2	Q5M8N6_MOUSE	Q5m8n6 mus musculus	630	52	77.6	741	2	Q8T9G4_DROME	Q60xr9 caenorhabdi
558	52	77.6	488	2	Q5RK12_RAT	Q5rk12 rattus norv	631	52	77.6	753	2	Q6H899_ORYSA	Q8t9g4 drosophila
559	52	77.6	488	2	Q6GQU7_MOUSE	Q6gqu7 mus musculus	632	52	77.6	757	2	Q9RV66_DEIRA	Q6h899 oryza sativ
560	52	77.6	490	2	Q69023_9GAMA	Q69023 human herpe	633	52	77.6	759	2	Q523F8_MAGGR	Q9rv66 deinococcus
561	52	77.6	493	2	Q9UVD1_PNECA	Q9uvd1 pneumocycati	634	52	77.6	773	2	Q7XHB6_ORYSA	Q523f8 magnaporthe
562	52	77.6	494	2	Q9FXA1_ARATH	Q9fxal arabidopsis	635	52	77.6	773	2	Q8S7Y7_ORYSA	Q7xhb6 oryza sativ
563	52	77.6	494	2	Q6ME19_FAROW	Q6me19 parachlamyd	636	52	77.6	774	2	Q86B84_CAEEL	Q8s7y7 oryza sativ
564	52	77.6	500	2	Q5QLD5_ORYSA	Q5qls5 oryza sativ	637	52	77.6	787	2	Q94096_PNECA	Q86b84 caenorhabdi
565	52	77.6	508	2	Q68Y49_ORYSA	Q68y49 oryza sativ	638	52	77.6	795	2	Q6EUQ1_ORYSA	Q94096 pneumocycati
566	52	77.6	508	2	Q9SLZ6_ORYSA	Q9slz6 oryza sativ	639	52	77.6	797	2	Q7XHB9_ORYSA	Q6euq1 oryza sativ
567	52	77.6	518	2	Q8T011_DROME	Q8t011 drosophila	640	52	77.6	797	2	Q8S7Z0_ORYSA	Q7xhb9 oryza sativ
568	52	77.6	518	2	Q9SE82_NAIZE	Q9seg2 zea mays (m	641	52	77.6	803	1	PDE4D_RAT	Q8s7z0 oryza sativ
569	52	77.6	519	2	Q48655_ORYSA	Q48655 oryza sativ	642	52	77.6	809	1	PDE4D_HUMAN	P14270 rattus norv
570	52	77.6	520	1	WASP_MOUSE	P730315 mus musculus	643	52	77.6	811	2	Q401R23_GIBZEA	Q8499 mus sapien
571	52	77.6	520	2	Q53WY0_MOUSE	Q53wy0 mus musculus	644	52	77.6	814	2	Q8C8U8_MOUSE	Q80t97 mus musculus
572	52	77.6	520	2	Q61078_MOUSE	Q61078 mus musculus	645	52	77.6	818	2	Q69375_9BETA	Q41r23 gibberella
573	52	77.6	526	2	Q5RGH5_BRARE	Q5rgw5 brachydanio	646	52	77.6	820	2	LOZEN_DROME	Q8c8u8 mus musculus
574	52	77.6	540	2	Q7XMB8_ORYSA	Q7xmb8 oryza sativ	647	52	77.6	826	1	Q8SNA3_MOUSE	Q69375 murid herpe
575	52	77.6	542	2	Q11448_NUCL	Q11448 anagrapha f	648	52	77.6	836	2	Q7S4M4_NEUCR	Q8sna3 mus musculus
576	52	77.6	542	2	Q92377_NPVBM	Q92377 bombyx mori	649	52	77.6	845	2	Q7S4M4_NEUCR	Q7s4m4 mus musculus
577	52	77.6	542	2	Q7WF9_NPVRO	Q7w9f9 rachiplusia	650	52	77.6	856	2	Q9LNA8_ARATH	Q9lna8 arabidopsis
578	52	77.6	543	1	VP61_NPVAC	Q3209 autographa	651	52	77.6	863	1	HCN2_MOUSE	Q8703 mus musculus
579	52	77.6	551	2	Q5ZB68_ORYSA	Q5zb68 oryza sativ	652	52	77.6	863	2	Q6BCT5_RAT	Q8703 mus musculus
580	52	77.6	560	2	Q9H985_HUMAN	Q9h985 homo sapien	653	52	77.6	864	2	Q8H837_ORYSA	Q6bct5 rattus norv
581	52	77.6	570	2	Q8L3T8_ORYSA	Q8l3t8 oryza sativ	654	52	77.6	870	2	Q8W028_ARATH	Q8h837 oryza sativ
582	52	77.6	576	2	Q742S0_MYCPA	Q742s0 mycobacteri	655	52	77.6	874	2	Q6AHU5_PNECA	Q8w028 arabidopsis
583	52	77.6	576	2	Q5ZK52_CHICK	Q5zk52 gallus gall	656	52	77.6	874	2	Q96VJ2_PNECA	Q6ahu5 pneumocycati
584	52	77.6	577	2	Q7XRW4_ORYSA	Q7xrw4 oryza sativ	657	52	77.6	893	2	Q6ZCX3_ORYSA	Q96vj2 pneumocycati
585	52	77.6	585	2	Q7XQW8_ORYSA	Q7xqw8 oryza sativ	658	52	77.6	902	2	Q6AHS3_PNECA	Q6zcx3 oryza sativ
586	52	77.6	588	2	Q9SNE9_ARATH	Q9sne9 arabidopsis	659	52	77.6	902	2	Q6AHV9_PNECA	Q6ahs3 pneumocycati
587	52	77.6	593	2	Q8LMC6_ORYSA	Q8lmc6 oryza sativ	660	52	77.6	909	2	Q8LIK9_ORYSA	Q6ahv9 pneumocycati
588	52	77.6	600	2	Q8N4J6_HUMAN	Q8n4j6 homo sapien	661	52	77.6	909	2	Q7XMC9_ORYSA	Q8lik9 oryza sativ
589	52	77.6	605	1	NCKX4_HUMAN	Q8nf22 homo sapien	662	52	77.6	909	2	Q56B20_RICBE	Q7xmc9 oryza sativ
590	52	77.6	605	1	NCKX4_MOUSE	Q8cgg8 mus musculus	663	52	77.6	910	2	Q5RIM8_MOUSE	Q56b20 rickettsia
591	52	77.6	613	2	Q6ZLD1_ORYSA	Q6zld1 oryza sativ	664	52	77.6	914	2	Q86YY0_HUMAN	Q5rim8 mus musculus
592	52	77.6	618	2	Q8S6K9_ORYSA	Q8s6k9 oryza sativ	665	52	77.6	927	2	Q7XSN2_ORYSA	Q86yy0 homo sapien
593	52	77.6	637	2	Q4P3H3_USTWA	Q4p3h3 ustilago ma	666	52	77.6	938	2	Q96V14_PNECA	Q7xsn2 oryza sativ
594	52	77.6	639	2	Q6DCV1_XENLA	Q6dcv1 xenopus lae	667	52	77.6	941	2	Q9C6N0_ARATH	Q96v14 pneumocycati
595	52	77.6	640	2	Q60593_HUMAN	Q60593 homo sapien	668	52	77.6	942	2	Q5YRU1_NOCFA	Q9c6n0 arabidopsis
596	52	77.6	643	2	Q4WCV2_ASPFU	Q4wcv2 aspergillus	669	52	77.6	947	2	Q6AHT0_PNECA	Q5yru1 nocardia fa
597	52	77.6	645	2	Q96EX0_HUMAN	Q96ex0 homo sapien	670	52	77.6	967	2	Q8TDL2_HUMAN	Q6aht0 pneumocycati
598	52	77.6	645	2	Q23374_ARATH	Q23374 arabidopsis	671	52	77.6	984	2	Q6L4M1_ORYSA	Q8tdl2 homo sapien
599	52	77.6	646	2	Q6PFT9_BRARE	Q6pft9 brachydanio	672	52	77.6	996	2	Q924M9_RAT	Q6l4m1 oryza sativ
600	52	77.6	647	2	Q4U2V7_CHLRE	Q4u2v7 chlamydomon	673	52	77.6	996	2	Q6JTV1_MOUSE	Q924m9 rattus norv
601	52	77.6	648	1	DREB_HUMAN	Q16643 homo sapien	674	52	77.6	996	2	Q6P730_RAT	Q6jtv1 mus musculus
602	52	77.6	651	2	Q9UFZ5_HUMAN	Q9ufz5 homo sapien	675	52	77.6	1010	2	Q8BUZ9_MOUSE	Q6p730 rattus norv
603	52	77.6	653	2	Q6CUH7_KUULA	Q6cuh7 kluyveromyc	676	52	77.6	1012	2	Q58H11_XENLA	Q8buz9 mus musculus
604	52	77.6	653	2	Q73TB8_MYCPA	Q73tb8 mycobacteri	677	52	77.6	1013	2	Q9NT81_HUMAN	Q58h11 xenopus lae
605	52	77.6	655	2	Q6P216_MOUSE	Q6p216 mus musculus	678	52	77.6	1015	2	Q8N3X1_HUMAN	Q9nt81 homo sapien
606	52	77.6	666	2	Q81Z27_HUMAN	Q81zz7 homo sapien	679	52	77.6	1021	2	Q9XUV2_CAEEL	Q8n3x1 homo sapien
607	52	77.6	666	2	Q60592_HUMAN	Q60592 homo sapien	680	52	77.6	1026	2	Q5VUR0_ORYSA	Q9xuv2 caenorhabdi
608	52	77.6	666	2	Q7ZSR6_HUMAN	Q7zsr6 homo sapien	681	52	77.6	1029	2	Q5SQ40_HUMAN	Q5vur0 oryza sativ
609	52	77.6	666	2	Q7TVG2_MYCBO	Q7tv92 mycobacteri	682	52	77.6	1031	2	Q5BNC8_MOUSE	Q5sq40 mus sapien
610	52	77.6	666	2	Q69740_MYCTU	Q69740 mycobacteri	683	52	77.6	1036	2	Q9C0C0_HUMAN	Q5bnc8 mus musculus
611	52	77.6	668	2	Q6Y8G4_RABIT	Q6y8g4 oryctolagus	684	52	77.6	1050	2	Q6Z003_MOUSE	Q9c0c0 homo sapien
612	52	77.6	670	2	Q8RVF5_ORYSA	Q8rvf5 oryza sativ	685	52	77.6	1052	2	Q9Y217_HUMAN	Q6z003 mus sapien
613	52	77.6	672	2	Q8LMH2_ORYSA	Q8lmh2 oryza sativ	686	52	77.6	1065	2	Q96SE1_HUMAN	Q9y217 homo sapien
614	52	77.6	684	2	Q9S282_STRCO	Q9s282 streptomyce	687	52	77.6	1077	2	Q5SQ42_HUMAN	Q6se1 mus musculus
615	52	77.6	687	2	Q948Y7_VOLVOX	Q948y7 volvox cart	688	52	77.6	1077	2	Q9JHC1_MOUSE	Q5sq42 homo sapien

689	52	77.6	1078	2	Q5SQ43_HUMAN	Q5sq43 homo sapien	762	52	77.6	2335	2	Q5RIM5_MOUSE	Q5rim5 mus musculus
690	52	77.6	1092	2	Q55JK0_CRYNE	Q55jk0 cryptococcu	763	52	77.6	2341	2	Q5RIM7_MOUSE	Q5rim7 mus musculus
691	52	77.6	1096	2	Q6MG49_RAT	Q6mg49 rattus norv	764	52	77.6	2343	2	Q7Z516_HUMAN	Q7z516 homo sapien
692	52	77.6	1098	2	Q9WTN8_RAT	Q9wtm8 rattus norv	765	52	77.6	2387	2	Q5PRE9_MOUSE	Q5pre9 mus musculus
693	52	77.6	1099	2	Q5KA01_CRYNE	Q5ka01 cryptococcu	766	52	77.6	2398	2	Q4Q3E5_LEIMA	Q4q3e5 leishmania
694	52	77.6	1126	2	Q5HYL9_HUMAN	Q5hyl9 homo sapien	767	52	77.6	2440	1	NCORI1_HUMAN	Q4q3e5 leishmania
695	52	77.6	1126	2	Q96SA6_HUMAN	Q96sa6 homo sapien	768	52	77.6	2453	1	NCORI1_MOUSE	Q75376 homo sapien
696	52	77.6	1126	2	Q9BCN4_HUMAN	Q9bcn4 homo sapien	769	52	77.6	2454	1	Q5RIM6_MOUSE	Q60974 mus musculus
697	52	77.6	1126	2	Q8IU41_DICDI	Q8iu41 dictyosteli	770	52	77.6	2601	2	Q4QGR7_LEIMA	Q4qgr7 leishmania
698	52	77.6	1126	2	Q54SP2_DICDI	Q54sp2 dictyosteli	771	52	77.6	2645	2	Q6XK19_DROME	Q6xk19 drosophila
699	52	77.6	1132	1	BAT3_HUMAN	P46379 homo sapien	772	52	77.6	2646	2	Q5CLM1_CRYHO	Q5clm1 cryptospori
700	52	77.6	1132	2	Q5STC1_HUMAN	Q5stc1 homo sapien	773	52	77.6	2706	2	Q15870_PLAFA	Q15870 plasmodium
701	52	77.6	1133	2	Q872C5_NEUCR	Q872c5 neurospora	774	52	77.6	3004	2	Q4SN41_TETNG	Q4sn41 tetradodon n
702	52	77.6	1144	2	Q7S3B2_NEUCR	Q7s3b2 neurospora	775	52	77.6	3110	1	HD_RAT	P51111 rattus norv
703	52	77.6	1154	2	Q9ZIR2_MOUSE	Q9zir2 mus musculus	776	52	77.6	3119	2	O8IHM0_PLAF7	O8ihm0 plasmodium
704	52	77.6	1159	2	Q69ZC2_MOUSE	Q69zc2 mus musculus	777	52	77.6	3139	2	Q9GM99_PIG	Q9gm99 sus scrofa
705	52	77.6	1160	2	Q5VMQ8_HUMAN	Q5vmq8 homo sapien	778	52	77.6	3795	2	Q4Q754_LEIMA	Q4q754 leishmania
706	52	77.6	1162	2	Q5SQ39_HUMAN	Q5sq39 homo sapien	779	51	76.1	57	2	Q13697_HUMAN	Q13697 homo sapien
707	52	77.6	1170	2	Q4IC26_GIBZE	Q4ic26 gibberella	780	51	76.1	70	2	Q9MAV4_ARATH	Q9ma4 arabadopsis
708	52	77.6	1171	2	Q94875_HUMAN	Q94875 homo sapien	781	51	76.1	88	2	Q4YAE9_PLABE	Q4yae9 plasmodium
709	52	77.6	1175	1	HCN4_RABIT	Q9tv66 oryctolagus	782	51	76.1	107	2	Q6VYS1_ORYSA	Q6vys1 oryza sativ
710	52	77.6	1175	2	Q4RBE3_TETNG	Q4reb3 tetradodon n	783	51	76.1	108	2	Q91AX2_9PAPI	Q91ax2 human papil
711	52	77.6	1190	2	Q54WB0_DICDI	Q54wb0 dictyosteli	784	51	76.1	127	2	Q8VAY0_WSSV	Q8vay0 white spot
712	52	77.6	1197	2	Q5W610_ORYSA	Q5w610 oryza sativ	785	51	76.1	134	2	Q9XIP3_ARATH	Q9xip3 arabadopsis
713	52	77.6	1205	2	Q4WGJ5_ASPFU	Q4wgj5 aspergillus	786	51	76.1	137	2	Q69U86_ORYSA	Q69u86 oryza sativ
714	52	77.6	1214	2	Q850T1_ORYSA	Q850t1 oryza sativ	787	51	76.1	138	2	Q84TS2_ORYSA	Q84ts2 oryza sativ
715	52	77.6	1220	1	VCIP1_MOUSE	Q8cdg3 mus musculus	788	51	76.1	151	2	Q7UE75_RHOBA	Q7ue75 rhodopirell
716	52	77.6	1221	1	VCIP1_RAT	Q8cf97 rattus norv	789	51	76.1	151	2	Q7UEG3_RHOBA	Q7ueg3 rhodopirell
717	52	77.6	1222	1	VCIP1_HUMAN	Q96jh7 homo sapien	790	51	76.1	153	2	Q67UD0_ORYSA	Q67ud0 oryza sativ
718	52	77.6	1222	2	Q504T4_HUMAN	Q504t4 homo sapien	791	51	76.1	157	2	Q7XV56_ORYSA	Q7xv56 oryza sativ
719	52	77.6	1229	2	Q95874_HUMAN	Q95874 homo sapien	792	51	76.1	175	2	Q6Z4V0_ORYSA	Q6z4v0 oryza sativ
720	52	77.6	1229	2	Q5SQ35_HUMAN	Q5sq35 homo sapien	793	51	76.1	182	2	Q86F22_SCHJA	Q86f22 schistosoma
721	52	77.6	1245	2	Q80T83_MOUSE	Q80t83 mus musculus	794	51	76.1	184	2	Q7XX24_ORYSA	Q7xx24 oryza sativ
722	52	77.6	1282	2	Q4Q6C2_LEIMA	Q4q6c2 leishmania	795	51	76.1	185	2	Q9XZ72_DEIRA	Q9xz72 deinococcus
723	52	77.6	1307	2	Q9LVN1_ARATH	Q9lvn1 arabadopsis	796	51	76.1	200	2	Q6YTD9_ORYSA	Q6ytd9 oryza sativ
724	52	77.6	1312	2	Q9NR59_HUMAN	Q9nr59 homo sapien	797	51	76.1	200	2	Q9CW04_MOUSE	Q9cw04 mus musculus
725	52	77.6	1312	2	Q7Z5V7_HUMAN	Q7z5v7 homo sapien	798	51	76.1	205	2	Q5CKE2_CRYHO	Q5cke2 cryptospori
726	52	77.6	1325	2	Q96PW1_HUMAN	Q96pw1 homo sapien	799	51	76.1	205	2	Q4NRN0_9DELT	Q4nrn0 anaeromykob
727	52	77.6	1343	2	Q9H7N4_HUMAN	Q9h7n4 homo sapien	800	51	76.1	208	2	Q61TA7_CABER	Q61ta7 caenorhabdi
728	52	77.6	1362	1	BRD4_HUMAN	Q60895 homo sapien	801	51	76.1	210	2	Q20327_CAEEL	Q20327 caenorhabdi
729	52	77.6	1384	2	Q4S827_TETNG	Q4s827 tetradodon n	802	51	76.1	215	2	Q5Z9C7_ORYSA	Q5z9c7 oryza sativ
730	52	77.6	1386	1	YLPML_MOUSE	Q9r017 mus musculus	803	51	76.1	221	2	Q49946_SOLTU	Q49946 solanum tub
731	52	77.6	1400	1	BRD4_MOUSE	Q9esu6 mus musculus	804	51	76.1	223	2	Q7XJ22_ORYSA	Q7xj22 oryza sativ
732	52	77.6	1449	2	Q4RY48_TETNG	Q4ry48 tetradodon n	805	51	76.1	234	2	Q5PXG5_EUPES	Q5pxg5 euphorbia e
733	52	77.6	1456	2	Q4NMZ8_9DELT	Q4nmz8 anaeromykob	806	51	76.1	243	2	Q6PI38_HUMAN	Q6pi38 homo sapien
734	52	77.6	1494	2	Q67FV1_HUMAN	Q67fv1 homo sapien	807	51	76.1	269	2	Q8ZYD1_PYRAE	Q8zyd1 pyrobaculum
735	52	77.6	1499	2	Q86UU0_HUMAN	Q86uu0 homo sapien	808	51	76.1	273	2	Q88AW2_PSESM	Q88aw2 pseudomonas
736	52	77.6	1542	1	SETBP_HUMAN	Q9y6x0 homo sapien	809	51	76.1	274	2	Q500G4_PSESY	Q500g4 pseudomonas
737	52	77.6	1556	2	Q4Q626_LEIMA	Q4q626 leishmania	810	51	76.1	277	2	Q9A5J8_CAUCR	Q9a5j8 caulobacter
738	52	77.6	1569	2	Q5CLH8_CRYHO	Q5clh8 cryptospori	811	51	76.1	278	1	TNPL6_RAT	P36940 rattus norv
739	52	77.6	1594	2	Q6PAL7_MOUSE	Q6pal7 mus musculus	812	51	76.1	281	1	ATF5_RAT	Q6p788 rattus norv
740	52	77.6	1595	2	Q5U5N7_MOUSE	Q5u5n7 mus musculus	813	51	76.1	283	1	ATF5_MOUSE	Q70191 mus musculus
741	52	77.6	1598	2	Q4REK0_TETNG	Q4reb0 tetradodon n	814	51	76.1	287	1	Q69X10_ORYSA	Q69x10 oryza sativ
742	52	77.6	1645	2	Q7TP73_RAT	Q7tp73 rattus norv	815	51	76.1	289	1	HXD8_MOUSE	P23463 mus musculus
743	52	77.6	1662	2	Q8X037_NEUCR	Q8x037 neurospora	816	51	76.1	290	2	Q9VC76_DROME	Q9vc76 drosophila
744	52	77.6	1662	2	Q7RWP2_NEUCR	Q7rwt2 neurospora	817	51	76.1	294	2	Q5VQC4_ORYSA	Q5vqc4 oryza sativ
745	52	77.6	1707	1	SET1_HUMAN	Q15047 homo sapien	818	51	76.1	306	2	P93845_PEA	P93845 pium sativ
746	52	77.6	1778	2	Q81WR5_HUMAN	Q81wr5 homo sapien	819	51	76.1	311	2	Q8LAX9_ARATH	Q8lax9 arabadopsis
747	52	77.6	1799	2	Q53LC9_ORYSA	Q53lc9 oryza sativ	820	51	76.1	311	2	Q9M9F0_ARATH	Q9m9f0 arabadopsis
748	52	77.6	1799	2	Q7XU11_ORYSA	Q7xu11 oryza sativ	821	51	76.1	312	2	Q8BZD7_MOUSE	Q8bzd7 mus musculus
749	52	77.6	1819	2	Q4TT23_9SPHN	Q4tt23 erythrobact	822	51	76.1	315	1	HAZ2D7_ARATH	Q8lme6 arabadopsis
750	52	77.6	1858	2	Q81ZD2_HUMAN	Q81zf2 homo sapien	823	51	76.1	318	2	Q23352_CAEEL	Q23352 caenorhabdi
751	52	77.6	1858	2	Q8NPF8_HUMAN	Q8nff8 homo sapien	824	51	76.1	319	2	Q6BQZ9_DEBHA	Q6bzq9 debaryomyce
752	52	77.6	1859	2	Q4PX2_USTWA	Q4px2 ustiliago ma	825	51	76.1	327	2	Q8BHG7_MOUSE	Q8bhg7 m mus muscu
753	52	77.6	1893	2	Q7ZWN3_XENLA	Q7zwn3 xenopus lae	826	51	76.1	328	2	Q5CHL3_CRYHO	Q5chl3 cryptospori
754	52	77.6	1938	2	Q66J90_XENLA	Q66j90 xenopus lae	827	51	76.1	338	2	Q6P0M8_HUMAN	Q6pom8 homo sapien
755	52	77.6	2008	2	Q5F3P8_CHICK	Q5f3p8 gallus gall	828	51	76.1	345	2	Q585B2_9TRYP	Q585b2 trypanosoma
756	52	77.6	2055	2	Q5NCQ1_MOUSE	Q5ncq1 mus musculus	829	51	76.1	347	2	Q9Y6U7_HUMAN	Q9y6u7 homo sapien
757	52	77.6	2123	2	Q9LPK3_ARATH	Q9lpk3 arabadopsis	830	51	76.1	348	1	Q8YJL9_BRUME	Q8yj19 brucella me
758	52	77.6	2135	2	Q4PFH2_USTWA	Q4pfh2 ustiliago ma	831	51	76.1	354	1	Q9XJ24_MOUSE	Q9xj24 mus musculus
759	52	77.6	2252	2	Q9VZX4_DROME	Q9vzx4 drosophila	832	51	76.1	362	1	GAW2_HUMAN	Q9nr24 homo sapien
760	52	77.6	2322	2	Q6UDW6_PLAFA	Q6udw6 plasmodium	833	51	76.1	374	2	Q66YD1_CHLRE	Q66y01 chlamydomon
761	52	77.6	2330	2	Q7SDG4_NEUCR	Q7sdg4 neurospora	834	51	76.1	378	2	Q61JL5_CABER	Q61j15 caenorhabdi

835	51	76.1	391	2	Q42532	ARATH	Q42532	arabidopsis	908	51	76.1	791	2	Q6ATQ3	ORYSA	Q6atq3	oryza sativ
836	51	76.1	391	2	Q60A53	METCA	Q60a53	methylococc	909	51	76.1	792	2	Q54B14	DICDI	Q54b14	dictyosteli
837	51	76.1	393	2	Q40552	TOBAC	Q40552	nicotiana t	910	51	76.1	796	2	Q8S6M0	ORYSA	Q8s6m0	oryza sativ
838	51	76.1	403	2	Q8BIH7	MOUSE	Q8biH7	mus musculu	911	51	76.1	807	2	Q5DTU6	MOUSE	Q5dtu6	mus musculu
839	51	76.1	412	2	Q5N8V9	ORYSA	Q5n8v9	oryza sativ	912	51	76.1	814	2	Q8LM34	ORYSA	Q8lm34	oryza sativ
840	51	76.1	413	2	Q9A4G2	CAUCR	Q9a4g2	caulobacter	913	51	76.1	814	2	Q9LY50	ARATH	Q9ly50	arabidopsis
841	51	76.1	415	2	Q873H2	NEUCR	Q873h2	neurospora	914	51	76.1	818	2	Q6K289	ORYSA	Q6k289	oryza sativ
842	51	76.1	415	2	Q29015	CEETA	Q29015	sus sp. pre	915	51	76.1	822	2	Q6L413	ORYSA	Q6l413	oryza sativ
843	51	76.1	418	2	Q60F43	ORYSA	Q60f43	oryza sativ	916	51	76.1	824	2	Q4RID3	TETNG	Q4rid3	tetradodon n
844	51	76.1	422	2	Q5Z513	ORYSA	Q5z513	oryza sativ	917	51	76.1	833	2	Q7XUV2	ORYSA	Q7xuv2	oryza sativ
845	51	76.1	429	2	Q4VXV5	HUMAN	Q4vxv5	homo sapien	918	51	76.1	835	2	Q9SDN5	TOBAC	Q9sdn5	nicotiana t
846	51	76.1	432	2	Q4VXV6	HUMAN	Q4vxv6	homo sapien	919	51	76.1	835	2	Q8H8K7	ORYSA	Q8h8k7	oryza sativ
847	51	76.1	433	2	Q4U2V8	CHLRE	Q4u2v8	chlamydomon	920	51	76.1	856	2	Q9VR54	DROME	Q9vr54	drosophila
848	51	76.1	439	2	Q5S808	CRYNE	Q5s808	cryptococcu	921	51	76.1	864	2	Q6Y444	ORYSA	Q6y444	oryza sativ
849	51	76.1	439	2	Q5K9G0	CRYNE	Q5k9g0	cryptococcu	922	51	76.1	881	2	Q6H7U3	ORYSA	Q6h7u3	oryza sativ
850	51	76.1	440	2	Q01672	ASCSU	Q01672	ascaris suu	923	51	76.1	918	2	Q4SKP19	TETNG	Q4skp19	tetradodon n
851	51	76.1	446	2	Q22458	CHLRE	Q22458	chlamydomon	924	51	76.1	978	2	Q4SXP7	TETNG	Q4sxp7	tetradodon n
852	51	76.1	448	1	WT1	RAT	P49952	rattus norv	925	51	76.1	1001	2	Q8VMD6	MOUSE	Q8vmd6	mus musculu
853	51	76.1	449	1	WT1	HUMAN	P19544	homo sapien	926	51	76.1	1006	2	Q9LMQ1	ARATH	Q9lmq1	arabidopsis
854	51	76.1	449	1	WT1	MOUSE	P22561	mus musculu	927	51	76.1	1027	2	Q9M3U2	DROME	Q9m3u2	drosophila
855	51	76.1	449	1	WT1	DIG	Q62651	sus scrofa	928	51	76.1	1067	2	Q6NSN0	BRARE	Q6nsn0	brachydanio
856	51	76.1	451	2	Q6EQW6	ORYSA	Q6eqw6	oryza sativ	929	51	76.1	1165	2	Q6ZHK2	ORYSA	Q6zhk2	oryza sativ
857	51	76.1	453	2	Q4VA17	MOUSE	Q4val17	mus musculu	930	51	76.1	1248	2	Q65589	ARATH	Q65589	arabidopsis
858	51	76.1	455	2	Q93XP5	ORYSA	Q93xp5	oryza sativ	931	51	76.1	1271	2	Q5FWL5	XENLA	Q5fwl5	xenopus lae
859	51	76.1	455	2	Q5Z514	ORYSA	Q5z514	oryza sativ	932	51	76.1	1293	2	Q6CAT2	YARLI	Q6cat2	varrowia li
860	51	76.1	460	2	Q5N906	ORYSA	Q5n906	oryza sativ	933	51	76.1	1392	2	Q9XER9	ARATH	Q9xer9	arabidopsis
861	51	76.1	484	2	Q5HZ98	XENLA	Q5hz98	xenopus lae	934	51	76.1	1422	2	Q53KD8	ORYSA	Q53kd8	oryza sativ
862	51	76.1	485	1	SSGP	VOLCA	P21997	volvox cart	935	51	76.1	1424	2	Q5XJV6	MOUSE	Q5xjv6	mus musculu
863	51	76.1	496	2	Q7QE74	ANOQA	Q7qe74	anopheles g	936	51	76.1	1456	2	Q52KF1	MOUSE	Q52kf1	mus musculu
864	51	76.1	497	1	WASP2	MOUSE	Q8bh43	mus musculu	937	51	76.1	1480	2	Q7SBI9	NEUCR	Q7sbi9	neurospora
865	51	76.1	497	2	Q9FP65	ORYSA	Q9fp66	oryza sativ	938	51	76.1	1486	2	Q54HS3	DICDI	Q54hs3	dictyosteli
866	51	76.1	497	2	Q5FWU0	RAT	Q5fwu0	rattus norv	939	51	76.1	1486	2	Q54HS3	DICDI	Q54hs3	dictyosteli
867	51	76.1	498	1	WASP2	HUMAN	Q9p6w5	homo sapien	940	51	76.1	1509	1	GSCRI	HUMAN	Q9nzml	homo sapien
868	51	76.1	501	2	Q69N33	ORYSA	Q69n33	oryza sativ	941	51	76.1	1568	2	Q5TKD9	ORYSA	Q5tkd9	oryza sativ
869	51	76.1	502	2	Q8NFC6	HUMAN	Q8nfc6	homo sapien	942	51	76.1	1689	2	Q5TQV0	ANOQA	Q5tqv0	anopheles g
870	51	76.1	503	2	Q661V2	XENLA	Q661v2	xenopus lae	943	51	76.1	1698	2	Q7QSB6	ANOQA	Q7qsb6	anopheles g
871	51	76.1	507	2	Q69XH2	ORYSA	Q69xh2	oryza sativ	944	51	76.1	1734	2	Q8MKW9	DROME	Q8mkw9	drosophila
872	51	76.1	507	2	Q8FY77	BRUSU	Q8fy77	brucella su	945	51	76.1	1743	2	Q966V0	DROME	Q966v0	drosophila
873	51	76.1	509	2	Q9GS27	HRIME	Q9gs27	hirudo medi	946	51	76.1	1814	2	Q4PBT2	USTWA	Q4pbt2	ustilago ma
874	51	76.1	516	2	Q4R808	ARATH	Q4r808	arabidopsi	947	51	76.1	1884	2	Q5CS67	CRYPV	Q5cs67	cryptospori
875	51	76.1	522	2	Q4R834	TETNG	Q4r834	tetradodon n	948	51	76.1	1931	2	Q4QX5	LEIMA	Q4q2x5	leishmania
876	51	76.1	524	2	Q9FN55	ARATH	Q9fn55	arabidopsis	949	51	76.1	1933	2	Q7EY18	ORYSA	Q7ey18	oryza sativ
877	51	76.1	528	2	Q89PV5	BRJIA	Q89pv5	bradyrhizob	950	51	76.1	1946	2	Q8N3R5	HUMAN	Q8n3r5	homo sapien
878	51	76.1	530	2	Q07701	CHV12	Q07701	cercopithe	951	51	76.1	1997	2	Q8LRM7	CHLRE	Q8lrm7	chlamydomon
879	51	76.1	535	1	Q51VQ7	MAGGR	Q51vq7	magnaporth	952	51	76.1	2162	2	Q55B12	DICDI	Q55e12	dictyosteli
880	51	76.1	555	1	GPI	CHLRE	Q9fpq6	chlamydomon	953	51	76.1	2274	2	Q54W13	DICDI	Q54w13	dictyosteli
881	51	76.1	567	2	Q6ID21	ANOQA	Q6id21	anopheles g	954	51	76.1	2376	2	Q966V1	DROME	Q966v1	drosophila
882	51	76.1	570	2	Q4TPN9	9SPHN	Q4tpn9	erythrobact	955	51	76.1	2376	2	Q9V5J0	DROME	Q9v5j0	drosophila
883	51	76.1	571	2	Q9RXG3	DEIRA	Q9rxg3	deinococcu	956	51	76.1	2386	1	EGFL4	HUMAN	Q7z7m0	homo sapien
884	51	76.1	584	2	Q82DU7	STRAW	Q82du7	streptomyce	957	51	76.1	2443	2	Q5CSU8	CRYPV	Q5csu8	cryptospori
885	51	76.1	599	2	P93797	VOLCA	P93797	volvox cart	958	51	76.1	2443	2	Q5CNM5	CRYHO	Q5cnm5	cryptospori
886	51	76.1	601	2	Q5BU15	FUGRU	Q5bul5	fugu rubrip	959	51	76.1	3251	2	Q8K4E0	MOUSE	Q8k4e0	mus musculu
887	51	76.1	606	2	Q852P0	VOLCA	Q852p0	volvox cart	960	51	76.1	3302	2	Q711X6	RAT	Q711x6	rattus norv
888	51	76.1	607	2	Q9LU79	ARATH	Q9lu79	arabidopsis	961	51	76.1	3345	2	Q5MJ67	HUMAN	Q5mj67	homo sapien
889	51	76.1	610	2	Q8GY16	ARATH	Q8gy16	arabidopsis	962	51	76.1	5085	1	PCLO	RAT	Q9v5j0	homo sapien
890	51	76.1	610	2	Q9MAX4	ARATH	Q9max4	arabidopsis	963	50.5	75.4	437	2	Q9STN0	ARATH	Q9stn0	arabidopsis
891	51	76.1	610	2	Q9SU56	ARATH	Q9su56	arabidopsis	964	50.5	75.4	639	2	Q55VN4	CRYNE	Q55vn4	cryptococcu
892	51	76.1	612	2	Q4SVK0	TETNG	Q4svk0	tetradodon n	965	50.5	75.4	639	2	Q5KKK5	CRYNE	Q5kkk5	cryptococcu
893	51	76.1	622	2	Q7TFD9	RHC6	Q7lfd9	rhesus cyto	966	50	74.6	20	2	Q7Z529	homo sapien	Q7z529	homo sapien
894	51	76.1	625	1	WASFA	HUMAN	Q8vmd6	mus musculu	967	50	74.6	33	2	Q02832	9PRIM	Q02832	gorilla gor
895	51	76.1	625	1	Q8S9B5	VOLCA	Q8s9b5	volvox cart	968	50	74.6	62	2	Q6Z053	ORYSA	Q6z053	oryza sativ
896	51	76.1	654	2	Q5ILT9	MAGGR	Q5ilt9	magnaporth	969	50	74.6	67	2	Q9Z118	RAT	Q9z118	rattus norv
897	51	76.1	675	2	Q8BI66	MOUSE	Q8bi66	mus musculu	970	50	74.6	67	2	Q75PL3	NUMME	Q75pl3	numida mele
898	51	76.1	684	2	Q7Y1D3	ORYSA	Q7y1d3	oryza sativ	971	50	74.6	76	2	Q8LQ30	ARATH	Q8lq30	arabidopsis
899	51	76.1	685	2	Q68A65	RAT	Q68a65	rattus norv	972	50	74.6	85	2	Q6YRL7	ORYSA	Q6yrl7	oryza sativ
900	51	76.1	701	2	Q7RWB1	NEUCR	Q7rwb1	neurospora	973	50	74.6	85	2	Q6ZLK9	ORYSA	Q6zlk9	oryza sativ
901	51	76.1	703	2	Q7S872	NEUCR	Q7s872	neurospora	974	50	74.6	89	2	Q6ZLN0	9PHVC	Q6zln0	etocarpus
902	51	76.1	712	2	Q5N7W0	ORYSA	Q5n7w0	oryza sativ	975	50	74.6	91	2	Q8ZT81	PYRAE	Q8zt81	pyrobaculum
903	51	76.1	715	1	AT12	HRVIF	P08314	human herpe	976	50	74.6	97	2	Q8MZM1	DROVI	Q8mzm1	drosophila
904	51	76.1	738	2	Q7PYV4	ANOQA	Q7pyv4	anopheles g	977	50	74.6	109	2	Q9BLB3	CAEEL	Q9blb3	caenorhabdi
905	51	76.1	759	2	Q8HXV1	RABIT	Q8hxv1	oryctolagus	978	50	74.6	117	2	Q5JUG1	THET8	Q5jg1	thermus the
906	51	76.1	760	2	Q9TOK5	ARATH	Q9tok5	arabidopsis	979	50	74.6	118	2	Q01614	PNECA	Q01614	pneumocysti
907	51	76.1	790	2	Q4KS29	9NUCL	Q4ks29	chrysodeixi	980	50	74.6	120	2	Q5A9M7	CANAL	Q5a9m7	candida alb

981 50 74.6 121 2 Q5A9C9 CANAL
 982 50 74.6 124 2 Q9FN57 ARATH
 983 50 74.6 126 2 Q8H604 ORYZA
 984 50 74.6 127 2 Q4THL7 TETNG
 985 50 74.6 128 2 Q6QNH4 ORTH
 986 50 74.6 134 2 Q5DGR5 SCHJA
 987 50 74.6 137 2 Q5SPH8 BRARE
 988 50 74.6 138 2 Q4UGR6 THEAN
 989 50 74.6 141 1 YPRO OWEFU
 990 50 74.6 142 2 Q00485 9CNID
 991 50 74.6 146 2 Q80VV6 MOUSE
 992 50 74.6 148 2 Q78DZ0 RATRT
 993 50 74.6 148 2 Q64371 RAT
 994 50 74.6 150 2 Q5D906 SCHJA
 995 50 74.6 150 2 Q6Z237 ORYZA
 996 50 74.6 157 2 Q7PN18 ANOGA
 997 50 74.6 158 2 Q6RW62 RAT
 998 50 74.6 161 2 Q23119 ARATH
 999 50 74.6 161 2 Q9M0N0 ARATH
 1000 50 74.6 162 2 Q23137 CAEBL

ALIGNMENTS

RESULT 1
 Q740X0 MYCPA
 ID Q740X0 MYCPA PRELIMINARY; PRT; 480 AA.
 AC Q740X0
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Hypothetical protein.
 GN OrderedLocNames=MAP122;
 OS Mycobacterium paratuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
 OC Mycobacterium avium complex (MAC).
 OX NCBI_TaxID=1770;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=K10;
 RA Li L., Bannantine J., Zhang Q., Amonsin A., Alt D., Kapur V.;
 RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE017231; AA030339.1; -; Genomic_DNA.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0016301; F:kinase activity; IEA.
 DR GO; GO:0000155; F:two-component sensor molecule activity; IEA.
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
 DR GO; GO:0007165; P:signal transduction; IEA.
 DR InterPro; IPR001594; ATPbind_ATPase.
 DR InterPro; IPR005467; His_kinase.
 DR InterPro; IPR003661; His_kinA_N.
 DR InterPro; IPR003660; His_kin_HAMP.
 DR InterPro; IPR004358; His_kin_like_C.
 DR Pfam; PF00672; HAMP; 1.
 DR Pfam; PF02518; HAMPase; 1.
 DR Pfam; PF00512; HisKA; 1.
 DR PRINTS; PR00344; BCTRUSENSOR.
 DR SMART; SM00304; HAMP; 1.
 DR SMART; SM00387; HATPase_c; 1.
 DR SMART; SM00388; HisKA; 1.
 DR PROSITE; PS50885; HAMP; 1.
 DR PROSITE; PS50109; HIS_KIN; 1.
 KW Complete proteome.
 SQ SEQUENCE 480 AA; 50782 MW; C85E333BEAD3181C CRC64;

Query Match 88.1%; Score 59; DB 2; Length 480;
 Best Local Similarity 90.0%; Pred. No. 82;
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PYPYPPPPIP 10

Db 121 PYPYPPPPPGP 130

RESULT 2

Q6Z1Z7 ORYZA
 ID Q6Z1Z7 ORYZA PRELIMINARY; PRT; 347 AA.
 AC Q6Z1Z7
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE PWP domain protein-like.
 GN Name=B1147B12.7-2;
 OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzoae; Oryza.
 OX NCBI_TaxID=39947;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Sasaki T., Matsumoto T., Katayose Y.;
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AP005406; BAD03547.1; -; Genomic_DNA.
 DR Gramene; Q6Z1Z7; -.
 SQ SEQUENCE 347 AA; 36953 MW; 3250DA6E418E70E8 CRC64;

Query Match 86.6%; Score 58; DB 2; Length 347;
 Best Local Similarity 80.0%; Pred. No. 75;
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PYPYPPPPIP 10
 Db 151 PYPYPPPPVP 160

RESULT 3

Q6ZTK4 HUMAN
 ID Q6ZTK4 HUMAN PRELIMINARY; PRT; 467 AA.
 AC Q6ZTK4
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Hypothetical protein FLJ44573.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
 OC Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Uterus;
 RA Tashiro H., Yamazaki M., Watanabe K., Kumagai A., Itakura S.,
 RA Fukuzumi Y., Fujimori Y., Komiyama M., Sugiyama T., Irie R.,
 RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
 RA Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
 RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Wagatsuma M.,
 RA Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima K., Sugiyama A.,
 RA Kawakami B., Suganuma Y., Sugano S., Nagahari K., Masuho Y., Nagai K.,
 RA Isogai T.;
 RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK126537; BAC86582.1; -; mRNA.
 SQ SEQUENCE 467 AA; 50372 MW; B27AB1A7A25421F0 CRC64;

Query Match 86.6%; Score 58; DB 2; Length 467;
 Best Local Similarity 80.0%; Pred. No. 1e+02;
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PYPYPPPPIP 10
 Db 176 PYPYPPPPVP 185

RESULT 4


```

Q6ZVU3_HUMAN
ID Q6ZVU3_HUMAN PRELIMINARY; PRT; 536 AA.
AC Q6ZVU3;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein FLJ42079.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE
RC TISSUE=Synovial membrane tissue;
RA Suzuki O., Sasaki N., Aotaka S., Shoji T., Ichihara T., Shiohata N.,
RA Matsumoto K., Hirano M., Sano S., Nomura R., Yoshikawa Y.,
RA Matsumura Y., Moriya S., Chiba E., Momiyama H., Onogawa S.,
RA Kaeriyama S., Satoh N., Matsunawa H., Takahashi E., Kataoka R.,
RA Kuga N., Kuroda A., Satoh I., Kamata K., Takami S., Terashima Y.,
RA Watanabe M., Sugiyama T., Irie R., Otsuki T., Sato H., Wakamatsu A.,
RA Iehli S., Yamamoto J., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T.,
RA Kimura K., Yamashita H., Matsuo K., Nakamura Y., Sekine M.,
RA Kikuchi H., Kanda K., Wagatsuma M., Murakawa K., Kanehori K.,
RA Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B., Suzuki Y.,
RA Sugano S., Nagahari K., Masuho Y., Nagai K., Isogai T.,
RA Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
RL EMBL; AK124073; BAC85767.1; -; mRNA.
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR InterPro; IPR000980; SH2.
DR Pfam; PF00017; SH2; 1.
DR ProDom; PD000093; SH2; 1.
DR SMART; SM00252; SH2; 1.
DR PROSITE; PS50001; SH2; 1.
SQ SEQUENCE 536 AA; 58716 MW; P6AF209F37091694 CRC64;

Query Match 86.6%; Score 58; DB 2; Length 536;
Best Local Similarity 80.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPYPYPPPPIP 10
Db 176 PPAYPPPPVP 185

RESULT 5
Q5ZH9 CHICK PRELIMINARY; PRT; 539 AA.
AC Q5ZH9;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein.
GN ORFNames=RCJMB04_3217;
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CB; TISSUE=Bursa;
RA Caldwell R.B., Kierzek A.M., Arakawa H., Bezzubov Y., Zaim J.,
RA Fiedler P., Kutter S., Blagodatski A., Kostovska D., Koter M.,
RA Plachy J., Carninci P., Hayaehizaki Y., Haverstedt J.M.;
RT "Full-length cDNAs from chicken bursa lymphocytes to facilitate
RT gene function analysis.";
RL Genome Biol. 6:R6-R6(2005).
CC -1- SIMILARITY: Contains 1 PH domain.
DR EMBL; AJ721025; CAG32684.1; -; mRNA.
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR InterPro; IPR001849; PH.
DR InterPro; IPR000980; SH2.

Q6ZVU3_MOUSE STANDARD; PRT; 559 AA.
AC Q6649;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE SH3 domain-binding protein 2 (3BP-2).
GN Name=Sh3bp2; Synonyms=3bp2;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridea; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=93174278; PubMed=8438166;
RA Ren R., Mayer B.J., Cicchetti P., Baltimore D.;
RT "Identification of a ten-amino acid proline-rich SH3 binding site.";
RL Science 259:1157-1161(1993).
CC -1- FUNCTION: Binds differentially to the SH3 domains of certain
CC proteins of signal transduction pathways. Binds to
CC phosphatidylinositols; linking the hemopoietic tyrosine kinase fcs
CC to the cytoplasmic membrane in a phosphorylation dependent
CC mechanism (By similarity).
CC -1- SIMILARITY: Contains 1 PH domain.
CC -1- SIMILARITY: Contains 1 SH2 domain.
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
DR EMBL; L14543; AAA37121.1; -; mRNA.
DR PIR; I49444; I49444.
DR HSP; Q9HB21; IBAZ.
DR Ensembl; ENSMUSG00000054520; Mus musculus.
DR MGI; MGI:1346349; Sh3bp2.
DR InterPro; IPR001849; PH.
DR InterPro; IPR011993; PH type.
DR InterPro; IPR000980; SH2.
DR Pfam; PF00169; PH; 1.
DR ProDom; PD000093; SH2; 1.
DR SMART; SM00233; PH; 1.
DR SMART; SM00252; SH2; 1.
DR PROSITE; PS50003; PH DOMAIN; 1.
DR PROSITE; PS50001; SH2; 1.
DR SH2 domain; SH3-binding.
KW DOMAIN 26 130 PH.
FT DOMAIN 455 553 SH2.
FT MOTIF 201 210 SH3-binding.
FT COMPBIAS 205 212 Poly-Pro.

```

```

FT COMPBIAS 236 240 Poly-Pro.
SQ SEQUENCE 559 AA; 63208 MW; EDFF1F11B259646E CRC64;

Query Match 86.6%; Score 58; DB 1; Length 559;
Best Local Similarity 80.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPPYPPPPPP 10
DB 201 PPAYPPPPVP 210

RESULT 7
ID Q5U3L0_MOUSE PRELIMINARY; PRT; 559 AA.
AC Q5U3L0;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DE SH3bp2 protein.
GN Name=Sh3bp2;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridea; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1];
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6; TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
Brownstein M.J., Udell T.B., Toshyuki S., Carninci P., Prange C.,
Raha S.S., Loquellano N.A., McKernan K.J., Casavant T.L., Scheetz T.E.,
Richardson S., McSwan P.J., Toshiyuki S., Carninci P., Prange C.,
Villalon D.K., Muzny D.M., Sodergren E.J., Malek J.A., Gunaratne P.H.,
Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
Butterfield Y.S., Schein J.E., Jones S.J.M., Marra M.A.;
"Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6; TISSUE=Brain;
RA Director MGC Project;
RL Submitted (NOV-2004) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 1 PH domain.
DR EMBL; BC085497; AAH85497.1; -; mRNA.
DR MGI; MGI:1346349; Sh3bp2.
DR GO; GO:0007442; P:intracellular signaling cascade; IEA.
DR InterPro; IPR001849; PH.
DR InterPro; IPR000980; SH2.
DR Pfam; PF00169; PH; 1.
DR Pfam; PF00017; SH2; 1.
DR ProDom; PD000093; SH2; 1.
DR SMART; SM00233; PH; 1.
DR SMART; SM00252; SH2; 1.
DR PROSITE; PS50003; PH DOMAIN; 1.
DR PROSITE; PS50001; SH2; 1.
DR SH2 domain.
SQ SEQUENCE 559 AA; 62188 MW; 62235DDE3A849E79 CRC64;

Query Match 86.6%; Score 58; DB 2; Length 559;
Best Local Similarity 80.0%; Pred. No. 1.2e+02;

```

```

Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPPYPPPPPP 10
DB 201 PPAYPPPPVP 210

RESULT 8
Q91252_MOUSE
ID Q91252_MOUSE PRELIMINARY; PRT; 559 AA.
AC Q91252;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE SH3-domain binding protein 2.
GN Name=Sh3bp2;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muroidae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1];
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=FVB/N;
RC TISSUE=Mammary tumor. Metallothionien-TGF alpha model. 10 month old
virgin mouse. Taken by biopsy.;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,
Raha S.S., Loquellano N.A., McKernan K.J., Abramson R.D., Mullaly S.J.,
Bosak S.A., McSwan P.J., Toshiyuki S., Carninci P., Prange C.,
Richardson S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
Butterfield Y.S., Schein J.E., Jones S.J.M., Marra M.A.;
"Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=FVB/N;
RC TISSUE=Mammary tumor. Metallothionien-TGF alpha model. 10 month old
virgin mouse. Taken by biopsy.;
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 1 PH domain.
DR EMBL; BC010198; AAH010198.1; -; mRNA.
DR Ensembl; ENSMUSG0000054520; Mus musculus.
DR MGI; MGI:1346349; Sh3bp2.
DR GO; GO:0007442; P:intracellular signaling cascade; IEA.
DR InterPro; IPR011993; PH type.
DR InterPro; IPR001849; PH.
DR InterPro; IPR000980; SH2.
DR Pfam; PF00169; PH; 1.
DR Pfam; PF00017; SH2; 1.
DR ProDom; PD000093; SH2; 1.
DR SMART; SM00233; PH; 1.
DR SMART; SM00252; SH2; 1.
DR PROSITE; PS50003; PH DOMAIN; 1.
DR PROSITE; PS50001; SH2; 1.
DR SH2 domain.
SQ SEQUENCE 559 AA; 62223 MW; FF485FDA78B40F20 CRC64;

Query Match 86.6%; Score 58; DB 2; Length 559;

```

Best Local Similarity 80.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPYPPPPPIP 10
Db 201 PPAYPPPPVP 210

RESULT 9
Q8C220_MOUSE PRELIMINARY; PRT; 559 AA.
ID Q8C220_MOUSE PRELIMINARY; PRT; 559 AA.
AC Q8C220;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Mus musculus B6-derived Cd11 +ve dendritic cells cDNA, RIKEN full-length enriched library, clone:F730032N16 product:SH3-domain binding protein 2, full insert sequence.
DE protein 2, full insert sequence.
GN Name=Sh3bp2;
OS Mus musculus (Mouse)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9; Carninci P., Hayashizaki Y.;
RA "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500; Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Iehii Y., Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S., Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I., Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R., Kadoya K., Matsuda H.A., Ashburner M., Batalov S., Casavant T., Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H., Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J., Schriml L.M., Straubli F., Suzuki R., Tomita M., Wagner L., Washio T., Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G., Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F., Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M., Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H., Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P., Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N., Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F., Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L., Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S., Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J;
RX The PANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100; Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M., Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [5]

RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600; Shibata K., Itoh M., Aizawa K., Nagao S., Sasaki N., Carninci P., Konno H., Akiyama J., Nishi K., Kitsuana T., Tashiro H., Itoh M., Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A., Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K., Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M., Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J., Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J;
RX Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P., Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W., Hayaehida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T., Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T., Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S., Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M., Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y., Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H., Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M., Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T., Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 1 PH domain.
DR EMBL; AK089455; BAC40891.1; -; mRNA.
DR Ensembl; ENSMUSG00000054520; Mus musculus.
DR MGI; MGI:1346349; Sh3bp2
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR InterPro; IPR001849; PH.
DR InterPro; IPR011993; PH type.
DR InterPro; IPR000980; SH2.
DR Pfam; PF00169; PH; 1.
DR Pfam; PF00017; SH2; 1.
DR ProDom; PD000093; SH2; 1.
DR SMART; SM00233; PH; 1.
DR SMART; SM00252; SH2; 1.
DR PROSITE; PS00003; PH DOMAIN; 1.
DR PROSITE; PS00001; SH2; 1.
KW SH2 domain.
SQ SEQUENCE 559 AA; 62218 MW; 62248F9AFB849E79 CRC64;
Query Match 86.6%; Score 58; DB 2; Length 559;
Best Local Similarity 80.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPYPPPPPIP 10
Db 201 PPAYPPPPVP 210

RESULT 10
3BP2_HUMAN STANDARD; PRT; 561 AA.
ID 3BP2_HUMAN STANDARD; PRT; 561 AA.
AC P78314; O00500; O15373; P78315;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE SH3 domain-binding protein 2 (3BP-2).
GN Name=SH3BP2; Synonyms=3BP2; ORFNames=RES4-23;
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM LONG).
RX MEDLINE=97446134; PubMed=929232; DOI=10.1006/geno.1997.4849; Bell S.M., Shaw M., Jou Y.-S., Myers R.M., Knowles M.A.;

RT "Identification and characterization of the human homologue of SH3BP2,
 RT an SH3 binding domain protein within a common region of deletion at
 RT 4p16.3 involved in bladder cancer.";
 RL Genomics 44:163-170(1997).
 RN [2]
 RP NUCLEOTIDE SEQUENCE [MRNA] (ISOFORMS LONG AND SHORT), AND TISSUE
 RP SPECIFICITY.
 RC TISSUE=Brain;
 RX MEDLINE=98403881; PubMed=9734812;
 RA Hadano S., Ishida Y., Ikeda J.-E.;
 RT "The primary structure and genomic organization of five novel
 RT transcripts located close to the Huntington's disease gene on human
 RT chromosome 4p16.3.";
 RL DNA Res. 5:177-186(1998).
 RN [3]
 RP NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM LONG).
 RC TISSUE=Tonsil;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Straubeberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shennen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T.J., Max S.I., Wang J., Hsieh F.,
 RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gumaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Feing J., Helton E., Kettaman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko V., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalka U., Smallus D.E.,
 RA Schnarch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [5]
 RP SPLICE ISOFORM(S) THAT ARE POTENTIAL NMD TARGET(S).
 RX PubMed=14759258; DOI=10.1186/gb-2004-5-2-r8;
 RA Hillman R.T., Green R.E., Brenner S.E.;
 RT "An unappreciated role for RNA surveillance.";
 RL Genome Biol. 5:RESEARCH008.1-RESEARCH008.16(2004).
 RN [6]
 RP VARIANTS CRBM GLN-415; PRO-415; ARG-418; HIS-418; LEU-418; ARG-420 AND
 RP GLU-420.
 RX MEDLINE=21275962; PubMed=11381256; DOI=10.1038/88832;
 RA Ueki Y., Tiziani V., Santenna C., Fukai N., Maulik C., Garfinkle J.,
 RA Ninomiya C., deAmoral C., Peters H., Habal M., Rhee-Morris L.,
 RA Dees J.B., Kreiborg S., Olsen B.R., Reichenberger E.;
 RT "Mutations in the gene encoding c-Abl-binding protein SH3BP2 cause
 RT cherubism.";
 RL Nat. Genet. 28:125-126(2001).
 RN [7]
 RP VARIANT CRBM ARG-420.
 RX PubMed=12908899; DOI=10.1002/ajmg.a.20226;
 RA Lo B., Faiyaz-Ul-Haque M., Kennedy S., Aviz R., Tsui L.C., Teebi A.S.;
 RT "Novel mutation in the gene encoding c-Abl-binding protein SH3BP2
 RT causes cherubism.";
 RL Am. J. Med. Genet. 121:37-40(2003).
 RN [8]
 RP VARIANT CRBM ARG-418.
 RX PubMed=14577811; DOI=10.1597/1545-1569(2003)040<0632:AMMITS>2.0.CO;2;
 RA Imai Y., Kanno K., Moriya T., Kayano S., Seino H., Matsubara Y.,

RA Yamada A.;
 RT "A missense mutation in the SH3BP2 gene on chromosome 4p16.3 found in
 RT a case of nonfamilial cherubism.";
 RL Cleft Palate Craniofac. J. 40:632-638(2003).
 CC -1- FUNCTION: Binds differentially to the SH3 domains of certain
 CC proteins of signal transduction pathways. Binds to
 CC phosphatidylinositols; linking the hemopoietic tyrosine kinase fcs
 CC to the cytoplasmic membrane in a phosphorylation dependent
 CC mechanism.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=Long;
 CC IsoId=P78314-1; Sequence=Displayed;
 CC Name=Short;
 CC IsoId=P78314-2; Sequence=VSP_004085, VSP_004086;
 CC Note=May be produced at very low levels due to a premature stop
 CC codon in the mRNA, leading to nonsense-mediated mRNA decay;
 CC -1- TISSUE SPECIFICITY: Expressed in a variety of tissues including;
 CC lung, liver, skeletal muscle, kidney and pancreas.
 CC -1- DISEASE: Defects in SH3BP2 are the cause of cherubism (CRBM)
 CC [MIM:118400]. CRBM is an autosomal dominant inherited syndrome
 CC characterized by excessive bone degradation of the upper and lower
 CC jaws, which often begins around three years of age. It is followed
 CC by development of fibrous tissue masses, which causes a
 CC characteristic facial swelling.
 CC -1- SIMILARITY: Contains 1 PH domain.
 CC -1- SIMILARITY: Contains 1 SH2 domain.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC -----
 CC EMBL; US6386; AAB72034.1; -; mRNA.
 CC EMBL; AB000462; BAA19119.1; -; mRNA.
 CC EMBL; AB000463; BAA19120.1; -; mRNA.
 CC EMBL; AF000936; AAB59973.1; -; mRNA.
 CC EMBL; BC022996; AAB22996.1; -; mRNA.
 CC HSSP; O08967; 1FHW.
 CC Ensembl; ENSG00000087266; Homo sapiens.
 CC HGNC; HGNC:10825; SH3BP2.
 CC H-InvDB; HIX0004037; -.
 CC MIM; 602104; -.
 CC MIM; 118400; -.
 CC GO; GO:0005070; F:SH3/SH2 adaptor activity; TAS.
 CC GO; GO:0007165; P:signal transduction; TAS.
 CC InterPro; IPR001849; PH.
 CC InterPro; IPR011993; PH_type.
 CC InterPro; IPR000980; SH2.
 CC Pfam; PF00169; PH; 1.
 CC Pfam; PF00017; SH2; 1.
 CC ProDom; PD000093; SH2; 1.
 CC SMART; SM00233; PH; 1.
 CC SMART; SM00252; SH2; 1.
 CC PROSITE; PS50003; PH DOMAIN; 1.
 CC PROSITE; PS50001; SH2; 1.
 CC Alternative splicing; Disease mutation; SH2 domain; SH3-binding.
 FT DOMAIN 26 130 PH.
 FT MOTIF 457 555 SH2.
 FT COMPBIAS 201 210 SH3-binding.
 FT COMPBIAS 205 212 Poly-Pro.
 FT COMPBIAS 236 240 Poly-Pro.
 FT VARSPLIC 81 97 VMEAAETTNNVPEFK -> QRPQPAQALSQTEAGP
 (in isoform Short).
 FT VARSPLIC 98 561 /FTId=VSP_004085.
 FT Missing (in isoform Short).
 FT VARIANT 415 415 /FTId=VSP_004086.
 FT R -> P (in CRBM).
 FT VARIANT 415 415 /FTId=VAR_013257.
 FT R -> Q (in CRBM).
 FT VARIANT 418 418 /FTId=VAR_013258.
 FT P -> H (in CRBM).

```

FT VARIANT 418 418 /FTid=VAR_013259.
FT FT P -> L (in CREM).
FT FT /FTid=VAR_013260.
FT VARIANT 418 418 P -> R (in CREM).
FT FT /FTid=VAR_013261.
FT VARIANT 420 420 G -> E (in CREM).
FT FT /FTid=VAR_013262.
FT VARIANT 420 420 G -> R (in CREM).
FT FT /FTid=VAR_013263.
FT CONFLICT 27 27 V -> L (in Ref. 3).
FT CONFLICT 224 224 H -> N (in Ref. 3).
FT CONFLICT 249 249 L -> R (in Ref. 3).
FT CONFLICT 251 251 A -> P (in Ref. 3).
SQ SEQUENCE 561 AA; 62244 MW; 69E6846A4F6D8F15 CRC64;

Query Match 86.6%; Score 58; DB 1; Length 561;
Best Local Similarity 80.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPYPPPPPIP 10
Db 201 PPAYPPPPVP 210

RESULT 11
Q7PW33_ANOGA PRELIMINARY; PRT; 1166 AA.
AC Q7PW33;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE ENSANGP0000016732 (Fragment).
GN ORFNames=ENSANGG0000014243;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Culicidae;
OC Anophelinae; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RT "Anopheles gambiae re-annotation.";
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
[2]
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic and membrane-associated (By
similarity).
CC -1- SIMILARITY: Contains 1 PH domain.
CC -1- CAUTION: The sequence shown here is derived from an
EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
preliminary data.
CC -1- SIMILARITY: Contains 1 PDZ (DHR) domain.
DR EMBL; AAB01008984; EAA15086.2; -; Genomic_DNA.
DR GO; GO:0005515; F:protein binding; IEA.
DR InterPro; IPR001478; PDZ.
DR InterPro; IPR001849; PH.
DR InterPro; IPR001660; SAM.
DR Pfam; PF00595; PDZ; 1.
DR Pfam; PF00169; PH; 1.
DR Pfam; PF00536; SAM; 1; 1.
DR SMART; SM00228; PDZ; 1.
DR SMART; SM00233; PH; 1.
DR SMART; SM00454; SAM; 1.
DR PROSITE; PS50106; PDZ; 1.
DR PROSITE; PS50003; PH DOMAIN; 1.
DR PROSITE; PS50105; SAM DOMAIN; 1.
FT NON TER 1166 1166
SQ SEQUENCE 1166 AA; 128485 MW; AC823199FOAF1484 CRC64;

```

```

Query Match 86.6%; Score 58; DB 2; Length 1166;
Best Local Similarity 80.0%; Pred. No. 2.6e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPYPPPPPIP 10
Db 976 PPPLPPPPVP 985

RESULT 12
Q621Z8_ORYSA PRELIMINARY; PRT; 1428 AA.
AC Q621Z8;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Putative PWWP domain protein.
GN Name=B1147B12.7-1;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsumoto T., Katayose Y.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP005406; BAD03546.1; -; Genomic_DNA.
DR Gramene; Q621Z8; -.
DR InterPro; IPR000313; PWWP.
DR InterPro; IPR006569; RPR.
DR Pfam; PF00855; PWWP; 1.
DR SMART; SM00293; PWWP; 1.
DR SMART; SM00582; RPR; 1.
DR PROSITE; PS50812; PWWP; 1.
KW Nucleotide-binding.
SQ SEQUENCE 1428 AA; 155202 MW; 03FB9D3E6CA7FB65 CRC64;

Query Match 86.6%; Score 58; DB 2; Length 1428;
Best Local Similarity 80.0%; Pred. No. 3.2e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPYPPPPPIP 10
Db 1186 PPPIPPPPVP 1195

RESULT 13
Q00487_9CNID PRELIMINARY; PRT; 172 AA.
AC Q00487;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Mini-collagen (Fragment).
GN Name=N-COL 4;
OS Hydra sp.
OC Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroidea; Anthomedusae;
OC Hydridae; Hydra.
OX NCBI_TaxID=6086;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=SP1;
RX MEDLINE=92064646; PubMed=1955459; DOI=10.1083/jcb.115.4.1159;
RA Kurz B.M., Holstein T.W., Petri B.M., Engel J., David C.N.;
RT "Mini-collagens in hydra nematocytes.";
RL J. Cell Biol. 115:1159-1169(1991).
DR EMBL; X61048; CAA43382.1; -; mRNA.
DR GO; GO:0005737; Cytoplasm; IEA.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006817; P:phosphate transport; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.

```

```
DR InterPro; IPR008160; Collagen.
DR InterPro; IPR000976; Wilms_tumour.
DR Pfam; PF01391; Collagen; 1.
DR PRINTS; PR00049; WILMSTUMOUR.
KW Collagen.
FT NON_TER
SQ SEQUENCE 172 AA; 16626 MW; F83AAD4369CFEB4F CRC64;

Query Match 85.1%; Score 57; DB 2; Length 172;
Best Local Similarity 90.0%; Pred. No. 46;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPYPPPPPIP 10
Db 57 PPPPPPPPIP 66

RESULT 14
Q6AVF5_ORYSA PRELIMINARY; PRT; 222 AA.
AC Q6AVF5;
DT 25-OCT-2004 (T-EMBLrel. 28, Created)
DT 25-OCT-2004 (T-EMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (T-EMBLrel. 28, Last annotation update)
DE Expressed protein.
GN Names:OSJNB0058G04.9;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Buell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Jones K.M.,
RA Overton II L.L., Taitzin T., Kim M.M., Bera J.J., Jin S.S.;
RA Fadon D.W., Tallon L.J., Koo H., Zismann V., Helao J., Blunt S.,
RA Vanaken S.S., Riedmuller S.B., Utterback T.T., Feldblyum T.V.,
RA Yang Q.Q., Haas B.J., Suh B.B., Peterson J.J., Quackenbush J.,
RA White O., Salzberg S.L., Fraser C.M.;
RT "Oryza sativa chromosome 3 BAC OSJNB0058G04 genomic sequence.";
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Buell R.;
RL SubMITTED (AUG-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC103551; AA778807.1; -; Genomic_DNA.
DR Gramene; Q6AVF5; -.
DR InterPro; IPR008889; VQ.
DR Pfam; PF05678; VQ; 1.
SQ SEQUENCE 222 AA; 22873 MW; 08EFB6375A5F9E19 CRC64;

Query Match 85.1%; Score 57; DB 2; Length 222;
Best Local Similarity 80.0%; Pred. No. 60;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPYPPPPPIP 10
Db 7 PPPPPPPPIP 16

RESULT 15
Q5AX03_EMENI PRELIMINARY; PRT; 231 AA.
AC Q5AX03;
DT 10-MAY-2005 (T-EMBLrel. 30, Created)
DT 10-MAY-2005 (T-EMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (T-EMBLrel. 30, Last annotation update)
DE Hypothetical protein.
GN CAPNames=AN7177.2;
OS Aspergillus nidulans FGSC A4.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; Emericella.
OX NCBI_TaxID=227321;
```

```

[1]
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=FGSC A4;
RA Birren B., Nusbaum C., Abouelheil A., Allen N., Anderson S.,
RA Arachchi H.M., Barna N., Bastien V., Bloom T., Boguslavskiy I.,
RA Boukhalter B., Butler J., Calvo S.E., Camarata J., Chang J.,
RA Choepei Y., Collymore A., Cooke P., Corum B., DeArelano K.,
RA Diaz J.S., Dodge K., Dooly K., Dorris L., Elkins T., Engels R.,
RA Erickson J., Faro S., Ferreira P., FitzGerald M., Gage D., Galagan J.,
RA Gardyna S., Gnerre S., Graham L., Grand-Pierre N., Hafez N.,
RA Hagopian D., Hagos B., Hall J., Horton L., Hulme W., Iliev I.,
RA Jaffe D., Johnson R., Jones C., Kamal M., Kamat A., Karatas A.,
RA Kells C., Landers T., Levine R., Lindblad-Toh K., Liu G., Lui A.,
RA Ma L.-J., Mabbitt R., MacLean C., Macdonald P., Major J., Manning J.,
RA Matthews C., Mauceli E., McCarthy M., Meldrim J., Meneus L.,
RA Mihova T., Mlenga V., Murphy T., Naylor J., Nguyen C., Nicol R.,
RA Nielsen C.B., Norbu C., O'Connor T., O'Donnell P., O'Neill D.,
RA Oliver J., Peterson K., Phunkhang P., Pierre N., Purcell S.,
RA Rachupka A., Ramasamy U., Raymond C., Retta R., Rise C., Rogov P.,
RA Roman J., Schauer S., Schuback R., Seaman S., Severy P., Smirnov S.,
RA Smith C., Spencer B., Stange-Thomann N., Stojanovic N., Stubbs M.,
RA Talanas J., Teafaye S., Theodore J., Topham K., Travers M.,
RA Vassiliev H., Venkataraman V.S., Viel R., Vo A., Wang S., Wilson B.,
RA Wu X., Wyman D., Young G., Zainoun J., Zembek L., Zimmer A., Zody M.,
RA Lander E.;
RT "Genome Sequence of Aspergillus nidulans.";
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AACD01000122; EAA61429.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 231 AA; 25893 MW; CBF091B4D16D6941 CRC64;

Query Match 85.1%; Score 57; DB 2; Length 231;
Best Local Similarity 80.0%; Pred. No. 63;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPYPPPPPIP 10
Db 88 PPSYPPPPPIP 97

RESULT 16
ATP5_HUMAN STANDARD; PRT; 282 AA.
AC Q5Y2D1; Q9BSA1; Q9UNQ3;
DT 16-OCT-2001 (Rel. 40, Created)
DT 13-SEP-2005 (Rel. 48, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Cyclic-AMP-dependent transcription factor ATP-5 (Activating
DE transcription factor 5) (Transcription factor ATP5).
GN Name=ATP5; Synonyms=ATFX;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA], AND VARIANT PRO-121.
RC TISSUE=Brain;
RX MEDLINE=20558615; PubMed=11087824; DOI=10.1073/pnas.240452197;
RA White J.H., McIlhinney R.A.J., Wise A., Ciruela F., Chan W.-Y.,
RA Emson P.C., Billinton A., Marshall F.H.;
RT "The GABAB receptor interacts directly with the related transcription
RT factors CREB2 and ATFX.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:13967-13972(2000).
RN [2]
RP NUCLEOTIDE SEQUENCE [MRNA].
RC TISSUE=Hepatoclastoma;
RX PubMed=15221005; DOI=10.1038/sj.onc.1207782;
RA Yamada S., Ohira M., Horie M., Ando K., Takayasu H., Suzuki Y.,
RA Sugano S., Hirata T., Goto T., Matsunaga T., Hiyama E., Hayashi Y.,
```

RA Ando H., Suita S., Kaneko M., Sasaki F., Hashizume K., Ohnuma N.,
RA Nakagawa A.;
RT "expression profiling and differential screening between
RT hepatoblastomas and the corresponding normal livers: identification of
RT high expression of the PLK1 oncogene as a poor-prognostic indicator of
RT hepatoblastomas.";
RL Oncogene 23:5901-5911(2004).
RN [3]
RP NUCLEOTIDE SEQUENCE [MRNA], AND VARIANT PRO-121.
RA Kohroki J., Tanaka K.;
RT "cDNA clone encoding leucine-zipper protein.";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
RN [4]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC TISSUE=Placenta;
RX MEDLINE=223388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner K.H., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whitney M., Touchman J.W., Green E.D., Dickinson M.C.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [5]
RP NUCLEOTIDE SEQUENCE [MRNA] OF 161-282.
RX MEDLINE=99303793; PubMed=10373550;
RA Pati D., Meistrich M.L., Plon S.E.;
RT "Human Cdc34 and Rad6B ubiquitin-conjugating enzymes target repressors
RT of cyclic AMP-induced transcription for proteolysis.";
RL Mol. Cell. Biol. 19:5001-5013(1999).
RN [6]
RP INTERACTION WITH CCND3, FUNCTION, AND SUBCELLULAR LOCATION.
RX PubMed=15358120; DOI=10.1016/j.bbrc.2004.07.053;
RA Liu W., Jiang J., Shen X., Sun Q., Liu W., Shen H., Gu J.;
RT "Cyclin D3 interacts with human activating transcription factor 5 and
RT potentiates its transcription activity.";
RL Biochem. Biophys. Res. Commun. 321:954-960(2004).
CC -1- FUNCTION: Transcriptional activator which binds the cAMP response
CC element (CRE) (consensus: 5'-GTGACGT[AG][AG]-3'), a sequence
CC present in many viral and cellular promoters and blocks the
CC differentiation of neuroprogenitor cells into neurons. Its
CC transcriptional activity is enhanced by CCND3 and slightly
CC inhibited by CDK4.
CC -1- SUBUNIT: Binds DNA as a dimer. Interacts with PTP4A1/PRL-1 (By
CC similarity). Interacts with CCND3, but not with CCND1 or CCND2.
CC -1- SUBCELLULAR LOCATION: Nuclear and cytoplasmic.
CC -1- SIMILARITY: Belongs to the bZIP family.
CC -1- SIMILARITY: Contains 1 bZIP domain.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR ENBL; AF305687; AAG22558.1; -; mRNA.
DR ENBL; AF3073613; BAD38650.1; -; mRNA.
DR ENBL; AB021663; BAA78477.2; -; mRNA.
DR ENBL; BC005174; AAH05174.1; -; mRNA.
DR ENBL; AF101388; AAD28370.1; -; mRNA.

DR HSP: P18848; 1C16.
DR TRANSFAC; T04877; -
DR Ensembl; ENSG00000169136; Homo sapiens.
DR HMC; HGNC:790; ATF5.
DR MIM; 606398; -
DR GO; GO:0003702; F:RNA polymerase II transcription factor acti. .; TAS.
DR GO; GO:0003714; F:transcription corepressor activity; TAS.
DR GO; GO:0000074; P:regulation of cell cycle; TAS.
DR GO; GO:0006357; P:regulation of transcription from RNA polyme. .; TAS.
DR InterPro; IPR011616; bZIP_1.
DR InterPro; IPR004827; TF_bZIP.
DR Pfam; PF00170; bZIP_1; 1.
DR SMART; SM00338; BRLZ; 1.
DR PROSITE; PS50217; bZIP; 1.
DR PROSITE; PS00036; bZIP_BASIC; FALSE_NEG.
KW Activator; DNA-binding; Multigene family; Nuclear protein;
KW Polymorphism; Transcription; Transcription regulation.
FT DOMAIN 236 250 Leucine-zipper (Probable).
FT DMAP 210 230 Basic motif.
FT REGION 119 217 Interaction with PTP4A1 (By similarity).
FT COMPTAS 123 139 Poly-Pro.
FT COMPTAS 186 194 Poly-Pro.
FT VARIANT 121 121 /FTID=VAR.022786.
FT CONFLICT 161 163 LLA -> RHE (in Ref. 5).
SQ SEQUENCE 282 AA; 30690 MW; B6D4CB6F631655B1 CRC64;
Query Match 85.1%; Score 57; DB 1; Length 282;
Best Local Similarity 80.0%; Pred. No. 77;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Qy 1 PPPYPPPPIP 10
Db 127 PPPLPPPPPLP 136
RESULT 17
Q81W12_HUMAN
ID Q81W12_HUMAN PRELIMINARY; PRT; 434 AA.
AC Q81W12;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Transcription factor AP-2 epsilon (Activating enhancer binding protein
DE 2 epsilon).
GN Homo sapiens (Human).
OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Uterus;
RX MEDLINE=223388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner K.H., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whitney M., Touchman J.W., Green E.D., Dickinson M.C.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human

```
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RC NUCLEOTIDE SEQUENCE.
RP TISSUE-Uterus;
RA Director MGC Project;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC041175; AAH41175.1; -; mRNA.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR004979; TF_AP2.
DR Pfam; PF03299; TF_AP-2; 1.
DR PRINTS; PR01748; AP2TNSCPCT.
SQ SEQUENCE 434 AA; 45308 MW; 6209DD6D07B8872C CRC64;

Query Match 85.1%; Score 57; DB 2; Length 434;
Best Local Similarity 80.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPYPVPPPIP 10
DB 48 PPYPQPPLP 57

RESULT 18
Q6VUC0_HUMAN
ID Q6VUC0_HUMAN PRELIMINARY; PRT; 442 AA.
AC Q6VUC0;
DT 03-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Transcription factor AP-2 epsilon (Activating enhancer binding protein 2 epsilon).
DE 2 epsilon).
GN Name=TFAP2E; ORFNames=RPS-983H21.2-001;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Heath P.;
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY226454; AAQ91614.1; -; mRNA.
DR EMBL; AL157951; CA123520.1; -; Genomic DNA.
DR EMBL; AC004865; CA123520.1; JOINED; Genomic DNA.
DR Ensembl; ENSG00000116819; Homo sapiens.
DR HGNC; HGNC:30774; TFAP2E.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR004979; TF_AP2.
DR Pfam; PF03299; TF_AP-2; 1.
DR PRINTS; PR01748; AP2TNSCPCT.
SQ SEQUENCE 442 AA; 46212 MW; 823209DA7C0EC527 CRC64;

Query Match 85.1%; Score 57; DB 2; Length 442;
Best Local Similarity 80.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPYPVPPPIP 10
DB 56 PPYPQPPLP 65
```

```
RESULT 19
Q61WU8_CABER
ID Q61WU8_CABER PRELIMINARY; PRT; 521 AA.
AC Q61WU8;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein CBG04249.
GN Name=CBG04249;
OS Caenorhabditis briggsae.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6238;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RL The C. briggsae Sequencing Consortium;
RG Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CAAC01000017; CAB60609.1; -; Genomic DNA.
DR Hypothetical protein.
SQ SEQUENCE 521 AA; 59983 MW; 40D732BAD6BAAAC80 CRC64;

Query Match 85.1%; Score 57; DB 2; Length 521;
Best Local Similarity 90.0%; Pred. No. 1.5e+02;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPYPVPPPIP 10
DB 188 PPYPVPPPIP 197

RESULT 20
Q5SIV4_THET8
ID Q5SIV4_THET8 PRELIMINARY; PRT; 590 AA.
AC Q5SIV4;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE DNA primase.
GN OrderedLocustNames=TTTHA1260;
OS Thermus thermophilus (strain HB8 / ATCC 27634 / DSM 579).
OC Bacteria; Deinococcus-Thermus; Deinococci; Thermales; Thermaceae;
OC Thermus.
OX NCBI_TaxID=300852;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=HB8;
RA Masui R., Kurokawa K., Nakagawa N., Tokunaga F., Koyama Y.,
RA Shibata T., Oshima T., Yokoyama S., Yasunaga T., Kuramitsu S.;
RT "Complete genome sequence of Thermus thermophilus HB8."
RL Submitted (NOV-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF008226; BAD71083.1; -; Genomic DNA.
DR GO; GO:0005658; C:alpha DNA polymerase:primase complex; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0003896; F:DNA primase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006304; P:DNA modification; IEA.
DR GO; GO:0006260; P:DNA replication; IEA.
DR GO; GO:0006269; P:DNA replication, synthesis of RNA primer; IEA.
DR InterPro; IPR006295; DNA primase.
DR InterPro; IPR006171; Toprim dom.
DR InterPro; IPR006647; Toprim primase.
DR InterPro; IPR006154; Toprim sub.
DR InterPro; IPR002694; Znf_CHC2.
DR Pfam; PF01807; zf-CHC2; 1.
DR Pfam; PD002276; Toprim primase; 1.
DR ProDom; PD002988; Znf_CHC2; 1.
DR SMART; SM00493; TOPRIM; 1.
DR SMART; SM00400; Znf_CHCC; 1.
DR TIGRFAMs; TIGR01391; dnaG; 1.
```



```

KW Complete proteome.
SQ SEQUENCE 590 AA; 66233 MW; 9524CA35246A6827 CRC64;

Query Match      85.1%; Score 57; DB 2; Length 590;
Best Local Similarity 90.0%; Pred. No. 1.7e+02;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPPYPPPPPI 10
   ||| |||||
Db 427 PPPPPPPPI 436

RESULT 21
ENAH_MOUSE
ID _ENAH_MOUSE STANDARD; PRT; 802 AA.
AC Q03173; P70431; P70432; P70433;
DT 01-OCT-1996 (Rel. 34, Created)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DT 10-MAY-2005 (Rel. 47, Last sequence update)
DE Enabled protein homolog (NPC derived proline-rich protein 1) (NDPP-1).
GN Name=Enah; Synonyms=Mena, Ndppl;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridea; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE (ISOFORM 1).
RC TISSUE=Brain;
RX MEDLINE=93041923; PubMed=1420303; DOI=10.1016/0167-4781(92)90156-T;
RA Sazuka T., Tomooka Y., Kathju S., Ikawa Y., Noda M., Kumar S.;
RT "Identification of a developmentally regulated gene in the mouse
RT central nervous system which encodes a novel proline rich protein.";
RL Biochim. Biophys. Acta 1132:240-248(1992).
RN [2]
RP NUCLEOTIDE SEQUENCE (ISOFORMS 2; 3; 4 AND 5), FUNCTION, AND
RP SUBCELLULAR LOCATION.
RC TISSUE=Brain;
RX MEDLINE=97015079; PubMed=8861907; DOI=10.1016/S0092-8674(00)81341-0;
RA Gertler F.B., Niebuhr K., Reinhard M., Wehland J., Soriano P.;
RT "Mena, a relative of VASP and Drosophila Enabled, is implicated in the
RT control of microfilament dynamics.";
RL Cell 87:227-239(1996).
RN [3]
RP FUNCTION, AND SUBUNIT.
RX MEDLINE=99166867; PubMed=10069337; DOI=10.1016/S0896-6273(00)81092-2;
RA Lanier L.M., Gates M.A., Witke W., Menzies A.S., Wehman A.M.,
RA Macklis J.D., Kwiakowski D., Soriano P., Gertler F.B.;
RT "Mena is required for neurulation and commissure formation.";
RL Neuron 22:313-325(1999).
RN [4]
RP ALTERNATIVE SPLICING (ISOFORM 6), PHOSPHORYLATION SITE TYR-557, AND
RP INTERACTION WITH ABI1.
RX PubMed=12672821; DOI=10.1074/jbc.M301447200;
RA Tani K., Sato S., Sukezane T., Kojima H., Hirose H., Hanafusa H.,
RA Shishido T.;
RT "Abi1 interactor 1 promotes tyrosine 296 phosphorylation of mammalian
RT enabled (Mena) by c-Abi kinase.";
RL J. Biol. Chem. 278:21685-21692(2003).
CC -!- FUNCTION: May be involved in microfilament assembly and cell
CC motility. Induces the formation of F-actin rich outgrowths in
CC fibroblasts. Required for neurulation and commissure formation.
CC -!- SUBUNIT: Binds profilin. Interacts with ABI1 and ROBO4.
CC -!- SUBCELLULAR LOCATION: Localized to focal adhesions and, to a
CC lesser extent, leading edges and stress fibers.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=6;
CC Name=5; Synonyms=Mena+++;
CC IsoId=Q03173-1; Sequence=Displayed;
CC Name=1;
CC IsoId=Q03173-2; Sequence=VSP_007255, VSP_007256;
CC Name=2; Synonyms=Mena;
CC IsoId=Q03173-3; Sequence=VSP_007259, VSP_007260;

```

```

CC CC Name=3; Synonyms=Mena+;
CC CC IsoId=Q03173-4; Sequence=VSP_007259;
CC CC Name=4; Synonyms=Mena++;
CC CC IsoId=Q03173-5; Sequence=VSP_007257, VSP_007258;
CC CC Name=6; Synonyms=Mena(S);
CC CC IsoId=Q03173-6; Sequence=VSP_007259, VSP_007260, VSP_010565;
CC CC -!- TISSUE SPECIFICITY: In the heart and testis and less so in the
CC lung, skeletal muscle, kidney, pancreas and brain. Isoform 6 is
CC expressed in brain and spleen, and embryo.
CC CC -!- SIMILARITY: Contains 1 WH1 domain.
CC CC -----
CC CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC CC -----
CC DR EMBL; D10727; BAA01570.1; -; mRNA.
CC DR EMBL; U72520; AAC52863.1; -; mRNA.
CC DR EMBL; U72521; AAC52864.1; -; mRNA.
CC DR EMBL; U72522; AAC52865.1; -; mRNA.
CC DR EMBL; U72523; AAC52866.1; -; mRNA.
CC DR PIR; S27200; S27200.
CC DR PDB; 1EVH; X-ray; A=1-112.
CC DR Ensembl; ENSMUSG00000022995; Mus musculus.
CC DR MGI; MGI:108360; Enah.
CC DR GO; GO:0030175; C:filopodium; IDA.
CC DR GO; GO:0005925; C:focal adhesion; IDA.
CC DR GO; GO:0030027; C:lamellipodium; IDA.
CC DR GO; GO:0001725; C:stress fiber; IDA.
CC DR GO; GO:0005522; P:profilin binding; IDA.
CC DR GO; GO:0017124; P:SH3 domain binding; IDA.
CC DR GO; GO:0007015; P:actin filament organization; IDA.
CC DR GO; GO:0008154; P:actin polymerization and/or depolymerization; IDA.
CC DR GO; GO:0007411; P:axon guidance; IGI.
CC DR GO; GO:0006928; P:cell motility; IDA.
CC DR GO; GO:0001843; P:neural tube closure; IGI.
CC DR InterPro; IPR00697; EVH1.
CC DR InterPro; IPR011993; PH_type.
CC DR InterPro; IPR00156; Ran_BP1.
CC DR Pfam; PF00568; WH1; 1.
CC DR SMART; SM00160; RanBD; 1.
CC DR SMART; SM00461; WH1; 1.
CC DR PROSITE; PS50229; WH1; 1.
CC DR 3D-structure; Alternative splicing; Developmental protein;
CC Differentiation; Neurogenesis; Phosphorylation.
KW DOMAIN 1 111 WH1.
FT COMPBIAS 442 464 Poly-Pro.
FT COMPBIAS 542 552 Poly-Pro.
FT COMPBIAS 562 574 Poly-Pro.
FT COMPBIAS 578 589 Poly-Pro.
FT COMPBIAS 593 605 Poly-Pro.
FT MOD_RES 557 557 Phosphotyrosine (by ABL).
FT VARSPPLIC 1 412 Missing (in isoform 1).
FT VARSPPLIC 117 135 Missing (in isoform 2, isoform 3 and
FT isoform 6).
FT VARSPPLIC 117 131 /FTId=VSP_007259.
FT VARSPPLIC 132 135 Missing (in isoform 4).
FT VARSPPLIC 259 500 C1FC -> VFYL (in isoform 4).
FT VARSPPLIC 535 631 Missing (in isoform 2 and isoform 6).
FT /FTId=VSP_007260.
FT POGVLGPPAPPPLPPSPGAYASALPPPPPPPPPLPS
FT TGPAPPPLPPPLNQAPPPLPPAPPLPASGIFSGSTSE
FT DNRELTGLAAAIAGA -> HRAIWDHLHLRHPLSHQALP
FT TPQHLLFPQDPLHHLCHPLVLLHLHPLHLKLLPLLP
FT HLLPLSPHLEFSLDPRQKTIAFNWTCSCNGS (in
FT isoform 1).
FT /FTId=VSP_007256.
FT VARSPPLIC 561 594 Missing (in isoform 6).
FT /FTId=VSP_010565.

```

```

SQ SEQUENCE 802 AA; 85844 MW; 592BB975EE20F77F CRC64;
Query Match 85.1%; Score 57; DB 1; Length 802;
Best Local Similarity 80.0%; Pred. No. 2.3e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPPYPPPPPIP 10
   ||| |||||
Db 454 PPLPLPPPLP 463

RESULT 22
O48682 ARATH PRELIMINARY; PRT; 820 AA.
AC O48682;
DT 01-JUN-1998 (TRMBLrel. 06, Created)
DT 01-JUN-1998 (TRMBLrel. 06, Last sequence update)
DT 01-OCT-2003 (TRMBLrel. 25, Last annotation update)
DE F316.8 protein.
GN Name=F316.8;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Federpiel N.A., Palm C.J., Conway A.B., Kurtz D.B., Conway A.R.,
RA Au M.J., Araujo R., Buehler E., Dewar K., Feng J., Kim C., Li Y.,
RA Yu G., Osborne B.I., Shinn P., Sun H., Toriumi M., Vysotskaia V.S.,
RA Yu G., Ecker J.J., Theologis A., Davis R.W.;
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC002396; AAC00575.1; -; Genomic_DNA.
DR PIR; T00645; T00645.
DR GO; GO:0003779; F-actin binding; IEA.
DR GO; GO:0016043; P-cell organization and biogenesis; IEA.
DR InterPro; IPR003104; FH2.
DR Pfam; PF02181; FH2; 1.
DR SMART; SM00498; FH2; 1.
SQ SEQUENCE 820 AA; 91171 MW; 770C7870CADEF965 CRC64;
Query Match 85.1%; Score 57; DB 2; Length 820;
Best Local Similarity 90.0%; Pred. No. 2.3e+02;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPPYPPPPPIP 10
   ||| |||||
Db 238 PPLPLPPPLP 247

RESULT 23
FRU_DROME STANDARD; PRT; 955 AA.
ID FRUINB1; O44708; P91618; P91619; Q24004; Q8IN80; Q8IN82; Q8IN83;
AC Q9GU18; Q9GU19; Q9GU20; Q9GU21; Q9GU22; Q9VE64; Q9VE65; Q9VE66;
AC Q9VE67;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Sex determination protein fruitless.
GN Name=fru; Synonyms=BTB-VI; ORFNames=CG14307;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP NUCLEOTIDE SEQUENCE (ISOFORMS MALE-I AND FEMALE-I), FUNCTION, TISSUE
RP SPECIFICITY, AND INTERACTION WITH TRA.
RC STRAIN=Oregon-R;
RX MEDLINE=97133215; PubMed=8978612; DOI=10.1016/S0092-8674(00)81802-4;
RA Ryner L.C., Goodwin S.F., Castrillon D.H., Anand A., Villeda A.,
RA Baker B.S., Hall J.C., Taylor B.J., Wasserman S.A.;

```

```

RT "Control of male sexual behavior and sexual orientation in Drosophila
RL by the fruitless gene.";
RN Cell 87:1079-1089(1996).
RP NUCLEOTIDE SEQUENCE (ISOFORM FEMALE-A), FUNCTION, AND TISSUE
RP SPECIFICITY.
RC STRAIN=Canton-S; TISSUE=Head;
RX MEDLINE=96382528; PubMed=8790392; DOI=10.1073/pnas.93.18.9687;
RA Ito H., Fujitani K., Usui K., Shimizu-Nishikawa K., Tanaka S.,
RA Yamamoto D.;
RT "Sexual orientation in Drosophila is altered by the satori mutation in
RT the sex-determination gene fruitless that encodes a zinc finger
RT protein with a BTB domain.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:9687-9692(1996).
RN [3]
RP NUCLEOTIDE SEQUENCE (ISOFORMS MALE-A; FEMALE-A; MALE-B; FEMALE-B;
RP TYPE-C; TYPE-D; MALE-E AND FEMALE-E), FUNCTION, TISSUE SPECIFICITY,
RP AND INTERACTION WITH TRA.
RC STRAIN=Canton-S; TISSUE=Head;
RX MEDLINE=20394292; PubMed=10934470; DOI=10.1038/35019537;
RA Uui-Aoki K., Ito H., Uui-Tei K., Takahashi K., Lukacsovich T.,
RA Avano W., Nakata H., Piao Z.F., Nilsson E.E., Tomida J.-Y.,
RA Yamamoto D.;
RT "Formation of the male-specific muscle in female Drosophila by ectopic
RT fruitless expression.";
RL Nat. Cell Biol. 2:500-506(2000).
RN [4]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Berkely;
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Anantides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bereman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA Foele C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacile J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Rainert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun B.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Wodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [5]
RP GENOME REANNOTATION, AND ALTERNATIVE SPLICING.
RX MEDLINE=22426089; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,

```

RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
RA Harris N.B., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.,
RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a
RT systematic review";
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
RN [6]
RP NUCLEOTIDE SEQUENCE OF 104-218.
RX MEDLINE=95024186; PubMed=7938017;
RA Zollman S., Godt D., Prive G.G., Couderc J.-L., Laski F.A.;
RT "The BTB domain, found primarily in zinc finger proteins, defines an
RT evolutionarily conserved family that includes several developmentally
RT regulated genes in *Drosophila*";
RL Proc. Natl. Acad. Sci. U.S.A. 91:10717-10721(1994).
RN [7]
RN FUNCTION.
RP MEDLINE=98078701; PubMed=9418892;
RX Dawalder B., Teujimoto S., Moss J., Mattox W.;
RA "The *Drosophila* taktout gene is regulated by the somatic sex-
RT determination pathway and affects male courtship behavior";
RL Genes Dev. 16:2879-2892(2002).
RN [8]
RP FUNCTION: Probably acts as a transcriptional regulator. Part of
CC the somatic sex determination hierarchy; sex determination genes
CC transform (tra) and transformer-2 (tra-2) switch fru splicing
CC from the male-specific pattern to the female-specific pattern
CC through activation of the female-specific fru 5' splice site.
CC Vital for the development of males and females. Controls the
CC development of the male specific abdominal muscle of Lawrence.
CC Plays a role in male courtship behavior and sexual orientation.
CC Enhances male-specific expression of takeout in brain-associated
CC fat body.
CC
CC -1- INTERACTION:
CC Q9V4B8:CG10323; NBExp=1; IntAct=EBI-196831, EBI-145217;
CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=10;
CC Comment=Sex specific splicing is controlled by tra and tra-2.
CC Sex-specific repression of the translation of fru mRNA might be
CC mediated by the binding of Tra to the fru mRNA in females;
CC Name=Male-A; Synonyms=E;
CC IsoId=Q8IN81-1; Sequence=Displayed;
CC Name=Female-A; Synonyms=C;
CC IsoId=Q8IN81-4; Sequence=VSP_050497;
CC Name=Female-B; Synonyms=F;
CC IsoId=Q8IN81-7; Sequence=VSP_050497, VSP_050501, VSP_050505;
CC Note=Ref.3 (AAG28589) sequence is in conflict in position:
CC 650:T->A;
CC Name=Female-E; Synonyms=H;
CC IsoId=Q8IN81-10; Sequence=VSP_050497, VSP_050500, VSP_050506;
CC Name=Female-I;
CC IsoId=Q8IN81-3; Sequence=VSP_050497, VSP_050498, VSP_050501,
CC VSP_050505;
CC Name=Male-B; Synonyms=B;
CC IsoId=Q8IN81-6; Sequence=VSP_050501, VSP_050505;
CC Name=Male-E; Synonyms=G;
CC IsoId=Q8IN81-12; Sequence=VSP_050500, VSP_050506;
CC Name=Male-I;
CC IsoId=Q8IN81-2; Sequence=VSP_050498, VSP_050501, VSP_050505;
CC Name=Type-C; Synonyms=D;
CC IsoId=Q8IN81-8; Sequence=VSP_050497, VSP_050499, VSP_050504;
CC Note=Ref.3 (AAG28590) sequence is in conflict in positions:
CC 539:S->P, 617:H->R;
CC Name=Type-D; Synonyms=A;

CC IsoId=Q8IN81-9; Sequence=VSP_050497, VSP_050502, VSP_050503;
CC -1- TISSUE SPECIFICITY: Expressed in parts of the adult male brain
CC associated with the courtship song and steps of the male
CC courtship. Also expressed in the larval and pupal male mushroom
CC body and optic lobe. Expressed in pupal female optic lobe.
CC -1- MISCELLANEOUS: Mutant males exhibit bisexual behavior; they court
CC females but are behaviorally sterile so fail to mate and they
CC exhibit vigorous courtship with other fru mutant males.
CC -1- SIMILARITY: Contains 1 BTB (POZ) domain.
CC -1- SIMILARITY: Contains 1 C2H2-type zinc finger.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; AF039231; AAB96677.1; -; mRNA.
CC EMBL; U72492; AAB92662.1; -; mRNA.
CC EMBL; D84437; BAA12663.1; -; mRNA.
CC EMBL; D84438; BAA12664.1; -; mRNA.
CC EMBL; AF220176; AAG28587.1; -; mRNA.
CC EMBL; AF220177; AAG28588.1; -; mRNA.
CC EMBL; AF220178; AAG28589.1; -; mRNA.
CC EMBL; AF220179; AAG28590.1; -; mRNA.
CC EMBL; AF220180; AAG28591.1; -; mRNA.
CC EMBL; AF220181; AAG28592.1; -; mRNA.
CC EMBL; AE003722; AAF55562.2; -; Genomic DNA.
CC EMBL; AE003722; AAF55563.2; -; Genomic DNA.
CC EMBL; AE003722; AAF55564.2; -; Genomic DNA.
CC EMBL; AE003722; AAF55565.2; -; Genomic DNA.
CC EMBL; AE003722; AAN13774.1; -; Genomic DNA.
CC EMBL; AE003722; AAN13775.1; -; Genomic DNA.
CC -----
CC Query Match 85.1%; Score 57; DB 1; Length 955;
CC Best Local Similarity 80.0%; Pred. No. 2.7e+02;
CC Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Oy 1 PPPVPPPPPIP 10
Db ||| |||||
78 PPPLPPPPPLP 87
RESULT 24
ID Q96JH1 HUMAN PRELIMINARY; PRT; 1134 AA.
AC Q96JH1;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE KIAA1856 protein (Fragment).
GN Name=KIAA1856;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=21245130; PubMed=11347906;
RA Nagase T., Nakayama M., Nakajima D., Kikuno R., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XX.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro";
RL DNA Res. 8:85-95(2001).
CC EMBL; AB058759; BAB47485.1; -; mRNA.
CC Ensembl; ENSG0000182095; Homo sapiens.
CC InterPro; IPR000637; A+T_hock.
CC PRINTS; PR00929; ATHOOK.
CC NON TER
CC SEQUENCE 1134 AA; 122299 MW; 1F1AA902C19F2EC2 CRC64;

```
Query Match      85.1%; Score 57; DB 2; Length 1134;
Best Local Similarity 80.0%; Pred. No. 3.2e+02;
Matches      8; Conservative      1; Mismatches      1; Indels      0; Gaps      0;

QY      1 PPPYPPPPPIP 10
Db      969 PPPLPPPPPLP 978
      |||||
      |||||

RESULT 25
Q6PLP6_CHLRE
ID      Q6PLP6_CHLRE PRELIMINARY;      PRT;      1146 AA.
AC      Q6PLP6;
DT      05-JUL-2004 (TrEMBLrel. 27, Created)
DT      05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT      05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE      Cell wall protein GP2 (Fragment).
GN      Name:GP2;
OS      Chlamydomonas reinhardtii.
OC      Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae;
OC      Chlamydomonadales; Chlamydomonadaceae; Chlamydomonas.
OC      NCBI_TaxID=3055;
RN      [1]
RP      NUCLEOTIDE SEQUENCE.
RC      STRAIN=C9;
RA      Ferris P.J., Small L., Goodenough U.W.;
RL      Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
DR      EMBL; AY596305; AAT02521.1; -; mRNA.
DR      InterPro; IPR002951; Atrophin.
DR      PRINTS; PRO1222; ATROPHIN.
FT      NON TER      1
SQ      SEQUENCE      1146 AA; 120311 MW; 749663DA78E316E3 CRC64;

Query Match      85.1%; Score 57; DB 2; Length 1146;
Best Local Similarity 80.0%; Pred. No. 3.3e+02;
Matches      8; Conservative      1; Mismatches      1; Indels      0; Gaps      0;

QY      1 PPPYPPPPPIP 10
Db      909 PPSPPPPPVP 918
      |||||
      |||||

RESULT 26
Q4WGB0_ASPPU
ID      Q4WGB0_ASPPU PRELIMINARY;      PRT;      1273 AA.
AC      Q4WGB0;
DT      13-SEP-2005 (TrEMBLrel. 31, Created)
DT      13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT      13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE      Hypothetical protein.
GN      ORFNames=Afu7G04400;
OS      Aspergillus fumigatus Af293.
OC      Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC      Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OC      NCBI_TaxID=330879;
RN      [1]
RP      NUCLEOTIDE SEQUENCE.
RC      STRAIN=Af293;
RA      Nerman W., Pain A., Anderson M.J., Wortman J., Kim H.Stanley.,
RA      Arroya J., Berriman M., Abe K., Archer D.B., Bermejo C., Bennett J.,
RA      Bowyer P., Chen D., Collins M., Coulson R., Davies R., Dyer P.S.,
RA      Farman M., Fedorova N., Fedorova N., Feldblyum T.V., Fischer R.,
RA      Foerker N., Fraser A., Garcia J.L., Garcia M.J., Goble A.,
RA      Goldman G.H., Gomi K., Griffith-Jones S., Gwilliam R., Haas B.,
RA      Haas H., Harris D., Horiuchi H., Huang J., Humphrey S., Jinenes J.,
RA      Kaller N., Khouri H., Kitamoto K., Kobayashi T., Kulkarni R.,
RA      Kumagai T., Lafont A., Latge J.-P., Li W., Lord A., Lu C.,
RA      Majores W.H., May G.S., Miller B.L., Mohamoud Y., Molina M., Monod M.,
RA      Mouyna I., Mulligan S., Murphy L., O'Neil S., Paulsen I.,
RA      Penalva M.A., Pertea S., Price C., Pritchard B.L., Quail M.A.,
RA      Rabinowitsch E., Rawlins N., Rajandream M.-A., Reichard U.,
RA      Renauld H., Robson G.D., Rodriguez de Cordoba S., Rodriguez-Pena J.M.,
RA      Ronning C.M., Rutter S., Salzberg S.L., Sanchez M.,
```

DR EMBL; AC117080; AAM45316.2; -: Genomic DNA.
DR EMBL; AAF10100030; EAL69520.1; -: Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 1324 AA; 148579 MW; F5C922EB79DDAB40 CRC64;

Query Match 85.1%; Score 57; DB 2; Length 1324;
Best Local Similarity 90.0%; Pred. No. 3.8e+02;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPYPPPPIP 10
||| |||||
Db 688 PPPPPPPPIP 697

RESULT 28

Q7SC01_NEUCR PRELIMINARY; PRT; 1395 AA.
AC Q7SC01;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Predicted protein (Hypothetical protein B22K18.130).
GN Name=NCU09482.1; Synonyms=B22K18.130;
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RC STRAIN=OR74A;
RA Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D., Jaffe D., FitzHugh W., Ma L.-J., Smirnov S., Purcell S., Reiman B., Elkins T., Engels S., Wang S., Nielsen C.B., Butler J., Endrizzi M., Qui D., Ianakiev P., Pedersen D., Nelson M., Washburne M., Selitrenikoff C.P., Kinsey J.A., Braun E.L., Zelter A., Schulte U., Kothe G.O., Jedd G., Mewes W., Staben C., Marcotte E., Greenberg D., Roy A., Foley K., Naylor J., Thomann N., Barrett R., Gnerre S., Kamal M., Kamysheles M., Maucelli E., Bielke C., Rudd S., Frishman D., Krystofova S., Rasmussen C., Metzenberg R.L., Perkins D.D., Kroken S., Cogoni C., Macino G., Catchside D., Li W., Pratt R.J., Osmani S.A., DeSouza C.C., Glass L., Orbach M.J., Berglund J., Voelker R., Varden O., Plamann M., Seiler S., Dunlap J., Radford A., Aramayo R., Natvig D.O., Alex L.A., Mannhaupt G., Ebbole D.J., Freitag M., Paulsen I., Sachs M.S., Lander E.S., Nussbaum C., Birren B.;
RT "The Genome Sequence of the Filamentous Fungus Neurospora crassa.";
RL Nature 0:0-0(2003).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Schulte U., Aign V., Hohseisel J., Brandt P., Fartmann B., Holland R., Nyakatura G., Mewes H.W., Mannhaupt G.;
RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RA German Neurospora genome project;
RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
CC -! CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AABX01000145; EAA33970.1; -: Genomic DNA.
DR EMBL; BX842597; CAE75735.1; -: Genomic_DNA.
DR HSPF; Q92831; IJM4.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0007049; P:cell cycle; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001487; Bromodomain.
DR PROSITE; PS50014; BROMODOMAIN_2; 1.
KW Hypothetical protein.
SQ SEQUENCE 1395 AA; 149956 MW; CB0D84322CDA8835 CRC64;

Query Match 85.1%; Score 57; DB 2; Length 1395;
Best Local Similarity 90.0%; Pred. No. 4e+02;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPYPPPPIP 10
||| |||||
Db 341 PPPPPPPPIP 350

RESULT 29

Q9VDD2_DROME PRELIMINARY; PRT; 1400 AA.
AC Q9VDD2; Q8SZ57;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE CG17299-PP, isoform F (UD22662p) (Loechnig isoform I)
GN Name=SNF4A-gamma; Synonyms=SNF4Agamma, loe; ORFNames=CG17299;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D., Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F., George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N., Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X., Chai B., Brannon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D., Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G., Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D., Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M., Beeson K.Y., Benos P.V., Bernier B.P., Bhandari D., Bolshakov S., Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P., Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I., Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P., de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M., Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P., Durbin K.J., Evangelista C.C., Ferraz C., Ferrieria S., Fleischmann W., Foeller C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K., Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M., Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J., Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C., Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A., Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z., Lasok P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X., Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D., Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A., Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L., Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M., Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G., Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H., Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T., Spier E., Spradling A.C., Stapleton M., Strong R., Sun E., Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X., Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J., Yao Q.A., Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A., Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L., Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O., Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA MEDLINE=22426065; PubMed=12537568;
RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A., Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A., George R.A., Hoskins R.A., Lavery T., Muzny D.M., Nelson C.R., Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E., Swirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C., Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT "Finishing a whole-genome shotgun: release 3 of the Drosophila melanogaster euchromatic genome sequence.";
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).

[3] NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426070; PubMed=12537573;
RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J.W., Svirekas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celinker S.E.;
RT "The transposable elements of the Drosophila melanogaster euchromatin:
a genomic perspective.";
RL Genome Biol. 3:RESEARCH0084.1-RESEARCH0084.20(2002).
[4]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426063; PubMed=12537572;
RA Miera S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Battencourt B.R., Celinker S.E., de Grey A.D.N.J., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Scapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
systematic review.";
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
[5]
RP NUCLEOTIDE SEQUENCE.
RX Berkeley Drosophila Genome Project;
RA Celinker S., Carlson J., Wan K., Pfeiffer B., Frise E., George R.,
RA Hoekins R., Stapleton M., Pacleb J., Park S., Svirekas R., Smith E.,
RA Yu C., Rubin G.;
RT "Drosophila melanogaster release 4 sequence.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
[6]
RP NUCLEOTIDE SEQUENCE.
RX FlyBase;
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
[7]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Barkley;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nuno J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celinker S.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
[8]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22344351; PubMed=12456644; DOI=10.1093/emboj/cdf636;
RA Tschape J.A., Hammarischmid C., Muhlig-Versen M., Athenstaedt K.,
RA Daum G., Kretschmar D.;
RT "The neurodegeneration mutant lochrig interferes with cholesterol
homeostasis and Appl processing.";
RL EMBO J. 21:6367-6376(2002).
[9]
RP NUCLEOTIDE SEQUENCE.
RA Tschape J.A.;
RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE003733; AAP55864.2; -; Genomic_DNA.
DR EMBL; AY070541; AAL48012.1; -; mRNA.
DR EMBL; AY166752; AAN85714.1; -; mRNA.
DR FlyBase; FBgn0025803; SNF4A-Gamma.
DR InterPro; IPR000644; CBS.
DR Pfam; PF00571; CBS; 2.
DR SMART; SM00116; CBS; 4.
SQ SEQUENCE 1400 AA; 152381 MW; 411B93CC6B9EC7AF CRC64;

Query Match 85.1%; Score 57; DB 2; Length 1400;
Best Local Similarity 90.0%; Pred. No. 4e+02;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPPPPPPIP 10
Db 280 PPPPPPPPIP 289

RESULT 30
Q752A6 ASHGO
ID Q752A6 ASHGO PRELIMINARY; PRT; 1918 AA.
AC Q752A6; 27, Created
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE AFR669WP.
GN Name=AFR669W;
OS Ashbya gossypii (Yeast) (Eremothecium gossypii).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Eremothecium.
OX NCBI_TaxID=33169;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=ATCC 10895;
RX PubMed=15001715; DOI=10.1126/science.1095781;
RA Dietrich F.S., Voegeli S., Brachat S., Lerch A., Gates K., Steiner S.,
RA Mohr C., Pohlmann R., Luedi P., Choi S., Wing R.A., Flavier A.,
RA Gaffney T.D., Philippsen P.;
RT "The Ashbya gossypii genome as a tool for mapping the ancient
Saccharomyces cerevisiae genome.";
RL Science 304:304-307(2004).
DR EMBL; AE016819; AAS54041.1; -; Genomic_DNA.
DR SMR; Q752A6; 1369-1776.
DR AGD; AFR669W; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003779; F:actin binding; IEA.
DR GO; GO:0017048; F:Rho GTPase binding; IEA.
DR GO; GO:0003700; F:actin cytoskeleton organization and biogenesis; IEA.
DR GO; GO:0003036; P:actin cytoskeleton organization and biogenesis; IEA.
DR GO; GO:0016043; P:cell organization and biogenesis; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR010472; Drf_FH3.
DR InterPro; IPR010473; Drf_GBD.
DR InterPro; IPR003104; FH2_actin_bd.
DR Pfam; PF06367; Drf_FH3; 1.
DR Pfam; PF06371; Drf_GBD; 1.
DR Pfam; PF02181; FH2; 1.
DR PRINTS; PR00049; WILMSTUMOUR.
DR SMART; SM00498; FH2; 1.
KW Complete proteome.
SQ SEQUENCE 1918 AA; 213018 MW; 2E07CC6AE7B9B0A6 CRC64;

Query Match 85.1%; Score 57; DB 2; Length 1918;
Best Local Similarity 90.0%; Pred. No. 5.6e+02;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPPPPPPIP 10
Db 1256 PPPPPPPPIP 1265

RESULT 31
Q5JM22 ORYSA
ID Q5JM22 ORYSA PRELIMINARY; PRT; 135 AA.
AC Q5JM22;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Hypothetical protein P0702H08.26.
GN Name=P0702H08.26;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoidae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsumoto T., Yamamoto K., Baba T., Katayose Y.,
RA Wu J., Niimura Y., Cheng Z., Nagamura Y., Antonio B.A., Kanamori H.,

```
RA Hosokawa S., Masukawa M., Arikawa K., Chiden Y., Hayashi M.,
RA Okamoto M., Ando T., Aoki H., Arita K., Hamada M., Harada C.,
RA Hijishita S., Honda M., Ichikawa Y., Ikonuma A., Iijima M., Ikeda M.,
RA Ikono M., Itoh T., Itoh Y., Iwabu A., Kibayashi N., Kamiya K.,
RA Kasawara W., Katagiri S., Kikuta A., Kobayashi T., Kono I.,
RA Machita K., Maehara T., Mizuno H., Mizubayashi T., Mukai Y.,
RA Nagasaki H., Nakashima M., Nakama Y., Nakamichi Y., Nakamura M.,
RA Namiki N., Negishi M., Ohta I., Ono N., Saij S., Sakai K., Shibata M.,
RA Shimokawa T., Shomura A., Song J., Takazaki Y., Terasawa K., Tsuji K.,
RA Waki K., Yamagata H., Yamane H., Yoshiki S., Yoshinara K., Yukawa K.,
RA Zhong H., Iwama H., Endo T., Ito H., Hahn J.H., Kim H.I., Eun M.Y.,
RA Yano M., Jiang J., Gojobori T.;
RT "The genome sequence and structure of rice chromosome 1.";
RL Nature 420:312-316(2002).
DR EMBL; AP003350; BAD87477.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 135 AA; 13727 MW; 72D1DF60F4B4E71 CRC64;

Query Match      83.6%; Score 56; DB 2; Length 135;
Best Local Similarity 80.0%; Pred. No. 46;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPYPPPPIP 10
Db 35 PPPPPPPPPV 44
|||||

RESULT 32
Q7YTN7 CAEBL
ID Q7YTN7 CAEBL PRELIMINARY; PRT; 170 AA.
AC Q7YTN7;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein F20G2.6a.
GN ORFNames=F20G2.6, F20G2.6a;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Bristol N2; PubMed=9851916;
RX MEDLINE=99069613; PubMed=9851916;
RG The C. elegans sequencing consortium;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL; Z79753; CAE17798.1; -; Genomic_DNA.
DR Ensembl; F20G2.6; Caenorhabditis elegans.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 170 AA; 18917 MW; 831A696032A97804 CRC64;

Query Match      83.6%; Score 56; DB 2; Length 170;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPYPPPP 8
Db 125 PPPYPPPP 132
|||||

RESULT 33
Q61FU6 CAEBR
ID Q61FU6 CAEBR PRELIMINARY; PRT; 173 AA.
AC Q61FU6;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein CBG11492.
GN Name=CBG11492;
OS Caenorhabditis briggsae.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
```

```
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6238;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RG The C. briggsae Sequencing Consortium;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CAAC0100054; CAB66248.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 173 AA; 19188 MW; B53564996FDBF2F4 CRC64;

Query Match      83.6%; Score 56; DB 2; Length 173;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPYPPPP 8
Db 128 PPPYPPPP 135
|||||

RESULT 34
Q52GY4 CAEBL
ID Q52GY4 CAEBL PRELIMINARY; PRT; 218 AA.
AC Q52GY4;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein F20G2.6b.
GN ORFNames=F20G2.6B;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Bristol N2; PubMed=9851916;
RX MEDLINE=99069613; PubMed=9851916;
RG The C. elegans sequencing consortium;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL; Z79753; CAI91165.1; -; Genomic_DNA.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 218 AA; 24544 MW; 9A9D014CAFC005DF CRC64;

Query Match      83.6%; Score 56; DB 2; Length 218;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPYPPPP 8
Db 173 PPPYPPPP 180
|||||

RESULT 35
Q64M77 ORYSA
ID Q64M77 ORYSA PRELIMINARY; PRT; 233 AA.
AC Q64M77;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein OSJNBa004005.28.
GN Name=OSJNBa004005.28;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsumoto T., Fujisawa M.;
```

RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 2, BAC
 RT clone:OSJNBa004C05."
 RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AP007253; BAD46728.1; -: Genomic_DNA.
 DR Gramene: Q64W77; -:
 DR InterPro: IPR001012; UBX.
 DR Pfam: PF00789; UBX; 1.
 DR PROSITE: PS50033; UBX; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 233 AA; 25299 MW; 350F3119708A4D7B CRC64;

Query Match 83.6%; Score 56; DB 2; Length 233;
 Best Local Similarity 80.0%; Pred. No. 81;
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPYPVPPPPIP 10
 ||||| |||||:
 Db 82 PPPPPPPPPVP 91

RESULT 36

Q4P543 USTMA PRELIMINARY; PRT; 247 AA.
 ID Q4P543
 AC Q4P543
 DT 13-SEP-2005 (TrEMBLrel. 31, Created)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
 DE Hypothetical protein.
 GN ORFNames=OM4770.1;
 OS Ustilago maydis 521.
 OC Eukaryota; Fungi; Basidiomycota; Ustilaginomycetes;
 OC Ustilaginomycetidae; Ustilaginales; Ustilaginaceae; Ustilago.
 OX NCBI_TaxID=237631;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=521;
 RA Birren B., Nusbaum C., Abebe A., Abouelleil A., Adekoya E.,
 RA Alt-zahra M., Allen N., Allen T., An P., Anderson M., Anderson S.,
 RA Arachchi H., Armbruster J., Rachtang P., Baldwin J., Barry A.,
 RA Bayul T., Blitshteyn B., Bloom T., Blye J., Boguslavskiy L.,
 RA Bolowsky M., Boukhgalter B., Brunache A., Butler J., Calixte N.,
 RA Carow S., Camarata J., Campo K., Chang J., Cheshatsang Y., Citroen M.,
 RA Collinmore A., Considine T., Cook A., Cooke P., Corum B., Cuomo C.,
 RA David R., Dawoe T., Degray S., Dodge S., Dooley K., Dorje P.,
 RA Dorjee K., Dorris L., Duffey N., Dupes A., Elkins T., Engels R.,
 RA Erickson J., Farina A., Faro S., Ferreira P., Fischer H.,
 RA Fitzgerald M., Foley K., Gage D., Galagan J., Gearin G., Gnerre S.,
 RA Gnirke A., Goyette A., Graham J., Grandbois E., Gyaltsen K., Hafez N.,
 RA Hagopian D., Hagos B., Hall J., Hatcher B., Heller A., Higgins H.,
 RA Honan T., Horn A., Houde N., Hughes L., Hulme W., Hugby E., Iliev I.,
 RA Jaffe D., Jones C., Kamal M., Kanat A., Kanvasselis M., Karlsson E.,
 RA Kalls C., Kieu A., Kisher P., Kodira C., Kulbokas E., Labutti K.,
 RA Lama D., Landers T., Leger J., Levine S., Lewis D., Lewis T.,
 RA Lindblad-toh K., Liu X., Lokitysang T., Lokitysang Y., Lucien O.,
 RA Lui A., Ma L.J., Mabbitt R., Macdonald J., Maclean C., Major J.,
 RA Manning J., Marabella R., Maru K., Matthews C., Mauceli E.,
 RA McCarthy M., McDonough S., McGhee T., Meldrim J., Meneus L.,
 RA Mesirov J., Mihalev A., Minova T., Mikkelsen T., Mlenga V., Moru K.,
 RA Mozes J., Mulrain L., Munson G., Naylor J., Neves C., Nguyen C.,
 RA Nguyen N., Nguyen T., Nicol R., Nielsen C., Nizzari M., Norbu C.,
 RA Norbu N., O'donnell P., Okoawo O., O'leary S., Omotohio B.,
 RA O'neill K., Osman S., Parker S., Perrin D., Phunkhang P., Piquani B.,
 RA Ratta R., Rachupka T., Ramasamy U., Rameau R., Ray V., Raymond C.,
 RA Rutman M., Schupbach R., Seanan C., Settippalli S., Sharpe P.,
 RA Sheridan J., Shupbach N., Shi J., Smirnov S., Smith C., Sougnez C.,
 RA Spencer B., Stalker J., Stange-thomann N., Stavropoulos S.,
 RA Stetson K., Stone C., Stone S., Stubbs M., Talamas J., Tchuinga P.,
 RA Tenzing P., Tesfaye S., Theodore J., Thoultsang Y., Topham K.,
 RA Towey S., Tsamla T., Tsomo N., Vallee D., Vassiliev H.,
 RA Vankataraman V., Vinson J., Vo A., Wade C., Wang S., Wangchuk T.,
 RA Wangdi T., Whittaker C., Wilkinson J., Wu Y., Wyman D., Yadav S.,
 RA Yang S., Yang X., Yeager S., Yee E., Young G., Zainoun J., Zembeck L.,

RA Zimmer A., Zody M., Lander E.;
 RT "The genome sequence of Ustilago maydis."
 RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
 CC -1- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 DR EMBL: AACP01000173; EAK86150.1; -: Genomic_DNA.
 KW Hypothetical protein.
 SQ SEQUENCE 247 AA; 27148 MW; 015A754D369D7EF2 CRC64;

Query Match 83.6%; Score 56; DB 2; Length 247;
 Best Local Similarity 80.0%; Pred. No. 86;
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPYPVPPPPIP 10
 ||||| |||||:
 Db 141 PPPPPPPPPVP 150

RESULT 37

Q5FVW5 XENTR
 ID Q5FVW5 XENTR PRELIMINARY; PRT; 353 AA.
 AC Q5FVW5
 DT 10-MAY-2005 (TrEMBLrel. 30, Created)
 DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
 DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
 DE LOC548379 protein (Fragment).
 GN Name=LOC548379;
 OS Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus; Silurana.
 OX NCBI_TaxID=8364;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Whole body;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Dege J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F.,
 RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Vallalona D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettaman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalek U., Smalls D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Whole body;
 RA Klein S., Garhard D.S.;
 RL Submitted (FEB-2005) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC089740; AA869740.1; -: mRNA.
 FT NON_TER
 SQ SEQUENCE 353 AA; 39442 MW; E731FF294870F9EC CRC64;

Query Match 83.6%; Score 56; DB 2; Length 353;
 Best Local Similarity 80.0%; Pred. No. 1.2e+02;
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPYPVPPPPIP 10
 ||||| |||||:
 Db 174 PPPPPPPPPVP 183


```
RESULT 38
EVL_RAT
ID EVL_RAT STANDARD; PRT; 393 AA.
AC O08719;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Ena/vasodilator stimulated phosphoprotein-like protein (Ena/VASP-like
DE protein).
DE
GN Name=Evl; Synonyms=Rnb6;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN (1)
RP NUCLEOTIDE SEQUENCE, TISSUE SPECIFICITY, AND DEVELOPMENTAL STAGE.
RC STRAIN=Wistar; TISSUE=Brain;
RX MEDLINE=97415794; PubMed=9268706; DOI=10.1006/bbrc.1997.7113;
RA Onca S., Mineta T., Kimoto M., Tabuchi K.;
RT "Differential display cloning of a novel rat cDNA (RNB6) that shows
RT high expression in the neonatal brain revealed a member of Ena/VASP
RT family."
RL Biochem. Biophys. Res. Commun. 237:307-312(1997).
CC -1- FUNCTION: Enhances actin nucleation and polymerization (By
CC similarity).
CC -1- SUBUNIT: Binds to the SH3 domains of ABL1, LYN and SRC. Also binds
CC to profilin and the WW domain of APBB1/FE65. Binds to SEMA6A (By
CC similarity).
CC -1- SUBCELLULAR LOCATION: Focal adhesions (By similarity).
CC -1- TISSUE SPECIFICITY: Expression detected in brain, spleen, thymus
CC and testis.
CC -1- DEVELOPMENTAL STAGE: In the brain, expression gradually increases
CC during embryonic development, reaches a maximum at postnatal day 1
CC and decreases thereafter.
CC -1- PTM: Phosphorylated by PKA; phosphorylation abolishes binding to
CC SH3 domains of ABL and SRC (By similarity).
CC -1- SIMILARITY: Contains 1 WH1 domain.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC ENBL; U70211; AAC53322.1; -; mRNA.
CC PIR; JC5614; JC5614.
CC HSP; P70429; 1QC6.
CC SMR; O08719; 1-113.
CC Ensembl; ENSRN000000014476; Rattus norvegicus.
CC RGD; 621150; Evl.
CC
CC GO; GO:0005737; Cytoplasm; ISS.
CC GO; GO:0005925; C:focal adhesion; ISS.
CC GO; GO:0030027; C:lamellipodium; ISS.
CC GO; GO:0005522; F:profilin binding; ISS.
CC GO; GO:0017124; F:SH3 domain binding; ISS.
CC GO; GO:0030048; P:actin filament-based movement; ISS.
CC GO; GO:0045010; P:actin nucleation; ISS.
CC GO; GO:0008154; P:actin polymerization and/or depolymerization; ISS.
CC GO; GO:0007411; P:axon guidance; NAS.
CC GO; GO:0051016; P:barbed-end actin filament capping; ISS.
CC GO; GO:0006928; P:cell motility; NAS.
CC GO; GO:0007166; P:cell surface receptor linked signal transdu. .; ISS.
CC GO; GO:0007417; P:central nervous system development; TAS.
CC GO; GO:0009887; P:organogenesis; ISS.
CC GO; GO:0030168; P:platelet activation; ISS.
CC InterPro; IPR000697; EVH1.
CC Pfam; PF00568; WH1.1.
CC SMART; SM00461; WH1.1.
CC PROSITE; PS0229; WH1; 1.

KW Phosphorylation; SH3-binding.
FT DOMAIN 1 112 WH1.
FT COMPBIAS 160 204 Pro-rich.
FT MOD_RES 327 327 Phosphoserine (By similarity).
FT MOD_RES 329 329 Phosphoserine (By similarity).
SQ SEQUENCE 393 AA; 42095 MW; 6371D91362925D4E CRC64;

Query Match 83.6%; Score 56; DB 1; Length 393;
Best Local Similarity 80.0%; Pred. No. 1.4e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPPPPPPP 10
Db 180 PPPPPPPPPV 189
|||||
|||||

RESULT 39
Q6PB99 MOUSE
ID Q6PB99 MOUSE PRELIMINARY; PRT; 393 AA.
AC Q6PB99;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Ena-vasodilator stimulated phosphoprotein.
DE
GN Name=Evl;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN (1)
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6; TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Klausner R.D., Feingold E.A., Grouse L.H., Derge J.G.,
RA Strausberg R.L., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN (2)
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6; TISSUE=Brain;
RX NIH MGC Project;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
DR ENBL; BC059810; AAH59810.1; -; mRNA.
DR SMR; Q6PB99; 1-113.
DR MGI; MGI:1194884; Evl.
DR GO; GO:0005515; F:protein binding; IPI.
DR InterPro; IPR000697; EVH1.
DR InterPro; IPR001960; WH1.
DR Pfam; PF00568; WH1; 1.
DR SMART; SM00461; WH1; 1.
SQ SEQUENCE 393 AA; 42093 MW; A76FD82A2B415D8D CRC64;

Query Match 83.6%; Score 56; DB 2; Length 393;
Best Local Similarity 80.0%; Pred. No. 1.4e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

Qy 1 PPPYPPPPIP 10
||| |||||:
Db 180 PPPPPPPVP 189

RESULT 40

ID Q501K3_MOUSE PRELIMINARY; PRT; 405 AA.
AC Q501K3;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DE Evi protein.
GN Mus musculus (Mouse).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6; TISSUE=Brain;
RX MEDLINE=2238825; Pubmed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Sapotnick M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Vallalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahney J., Helton E., Kettelman M., Madan A.C., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.; Smailus D.E.,
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6; TISSUE=Brain;
RG NIH MGC Project;
RL Submitted (MAY-2005) to the EMBL/GenBank/DBSJ databases.
DR EMBL; BC096017; AAH96017.1; -; mRNA.
DR SMR; Q501K3; 1-113.
DR MGI; MGI:1194884; Evi.
DR GO; GO:0005515; F:protein binding; IPI.
DR InterPro; IPR000697; EVH1.
DR InterPro; IPR011988; MHCII_invariant.
DR InterPro; IPR011993; PH_type.
DR Pfam; PF00568; WH1; 1.
DR SMART; SM00461; WH1; 1.
SQ SEQUENCE 405 AA; 43420 MW; 8D76F377139534EC CRC64;

Query Match 83.6%; Score 56; DB 2; Length 405;
Best Local Similarity 80.0%; Pred. No. 1.4e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPYPPPPIP 10
||| |||||:
Db 180 PPPPPPPVP 189

RESULT 41

Q9SBM1_VOLCA PRELIMINARY; PRT; 409 AA.
ID Q9SBM1;
AC Q9SBM1;

DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hydroxyproline-rich glycoprotein DZ-HRGP precursor.
GN Name=HRGP gene;
OS Volvox carteri f. nagariensis.
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae;
OC Chlamydomonadales; Volvocaceae; Volvox.
OX NCBI_TaxID=3068;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=HK10;
RX MEDLINE=20044763; Pubmed=10574980; DOI=10.1074/jbc.274.49.35023;
RA Ender F., Hallmann A., Amon P., Sumper M.;
RT "Response to the sexual pheromone and wounding in the green alga
RT volvox: induction of an extracellular glycoprotein consisting almost
RT exclusively of hydroxyproline.";
RL J. Biol. Chem. 274:35023-35028(1999).
DR EMBL; AJ242540; CAB62280.1; -; mRNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005199; F:structural constituent of cell wall; IEA.
DR GO; GO:0005249; F:voltage-gated potassium channel activity; IEA.
DR GO; GO:0006813; P:potassium ion transport; IEA.
DR InterPro; IPR005199; F:structural constituent of cell wall; IEA.
DR InterPro; IPR005404; KV3.3 channel.
DR InterPro; IPR003882; Pistil_extensin.
DR PRINTS; PR01222; ATROPHIN.
DR PRINTS; PR01582; KV33CHANNEL.
DR PRINTS; PR01218; PSTEXTENSIN.
KW Signal.
FT SIGNAL 1 18 Potential.
FT CHAIN 18 409 hydroxyproline-rich glycoprotein DZ-HRGP.
SQ SEQUENCE 409 AA; 41547 MW; CD0749C6AF02BD74 CRC64;

Query Match 83.6%; Score 56; DB 2; Length 409;
Best Local Similarity 80.0%; Pred. No. 1.4e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPYPPPPIP 10
||| |||||:
Db 87 PPPPPPPVP 96

RESULT 42
EVL_MOUSE
ID EVL_MOUSE STANDARD; PRT; 414 AA.
AC P70429; Q9ERU8;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Ena/vasodilator stimulated phosphoprotein-like protein (Ena/VASP-like
DE protein).
GN Name=Evl;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE (ISOFORM 1).
RC TISSUE=Brain;
RX MEDLINE=97015079; Pubmed=8861907; DOI=10.1016/S0092-8674(00)81341-0;
RA Gertler F.B., Niebuh K., Reinhard M., Weiland J., Soriano P.;
RT "Wena, a relative of VASP and Drosophila Enabled, is implicated in the
RT control of microfilament dynamics.";
RL Cell 87:227-239(1996).
RN [2]
RP NUCLEOTIDE SEQUENCE (ISOFORM 2), FUNCTION, SUBUNIT, SUBCELLULAR
RP LOCATION AND PHOSPHORYLATION BY PKA.
RC STRAIN=C57BL/6J;
RX Pubmed=10945997; DOI=10.1074/jbc.M006274200;
RA Lambrechts A., Kwiatkowski A.V., Lanier L.M., Bear J.E.,
RA Vandekerckhove J., Ampe C., Gertler F.B.;

RT "CAMP-dependent protein kinase phosphorylation of EVL, a Mena/VASP
RT relative, regulates its interaction with actin and SH3 domains."
RL J. Biol. Chem. 275:36143-36151(2000).
RN [3]
RP INTERACTION WITH SEMA6A.
RX PubMed=1093894; DOI=10.1074/jbc.M006316200;
RA Klostermann A., Lutz B., Gertler F., Behl C.;
RT "The orthologous human and murine semaphorin 6A-1 proteins (SEMA6A-
RT 1/SEMA6A-1) bind to the enabled/vasodilator-stimulated phosphoprotein-
RT like protein (EVL) via a novel carboxyl-terminal zyxin-like domain."
RL J. Biol. Chem. 275:39647-39653(2000).
CC -1- FUNCTION: Enhances actin nucleation and polymerization.
CC -1- SUBUNIT: Binds to the SH3 domains of ABL1, LYN and SRC. Also binds
CC to profilin and the WW domain of APEB1/PE65. Binds to SEMA6A.
CC -1- SUBCELLULAR LOCATION: Focal adhesions.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=2;
CC IsoId=P70429-1; Sequence=Displayed;
CC Name=1;
CC IsoId=P70429-2; Sequence=VSP_004045;
CC -1- PFM: Phosphorylated by PKA; phosphorylation abolishes binding to
CC SH3 domains of ABL and SRC.
CC -1- SIMILARITY: Contains 1 WH1 domain.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR ENBL; U72519; AAC52862.1; -; mRNA.
DR ENBL; AF279662; AAG23653.1; -; mRNA.
DR PDB; 1QC6; X-ray; A/B=1-130.
DR Ensembl; ENSMUSG00000021262; Mus musculus.
DR MGI; MGI:1194884; Evl.
DR GO; GO:0005737; Cytoplasm; IDA.
DR GO; GO:0005925; C:focal adhesion; IDA.
DR GO; GO:0030027; C:lamellipodium; IDA.
DR GO; GO:0005522; F:profilin binding; IDA.
DR GO; GO:001124; F:SH3 domain binding; IDA.
DR GO; GO:0030048; P:actin filament-based movement; NAS.
DR GO; GO:0045010; P:actin nucleation; NAS.
DR GO; GO:0008154; P:actin polymerization and/or depolymerization; IDA.
DR GO; GO:0007411; P:axon guidance; NAS.
DR GO; GO:0051016; P:barbed-end actin filament capping; NAS.
DR GO; GO:0007166; P:cell surface receptor linked signal transdu. .; ISS.
DR GO; GO:0007399; P:neurogenesis; ISS.
DR GO; GO:0030168; P:platelet activation; NAS.
DR InterPro; IPR000697; EVH1.
DR Pfam; PF00568; WH1; 1.
DR SMART; SM00461; WH1; 1.
DR PROSITE; PS50229; WH1; 1.
KW 3D-structure; Alternative splicing; Phosphorylation; SH3-binding.
FT DOMAIN 1 112 WH1.
FT COMPBTAS 160 204 Pro-rich.
FT MOD RES 327 327 Phosphoserine (By similarity).
FT MOD RES 329 329 Phosphoserine (By similarity).
FT VARSPLIC 341 361 Missing (in isoform 1).
FT /FTid=VSP_004045.
FT STRAND 4 17
FT TURN 18 21
FT STRAND 22 25
FT STRAND 34 41
FT TURN 42 45
FT STRAND 46 52
FT TURN 54 56
FT STRAND 59 64
FT TURN 67 68
FT STRAND 70 75
FT TURN 76 77
FT STRAND 78 82

FT STRAND 87 92
FT HELIX 95 111
FT TURN 112 112
SQ SEQUENCE 414 AA; 44337 MW; 146A018BCD6CA370 CRC64;
Query Match 83.6%; Score 56; DB 1; Length 414;
Best Local Similarity 80.0%; Pred. No. 1.5e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Qy 1 PPPYPPPPIP 10
Db 180 PPPPPPPPPV 189
|||||
RESULT 43
EVL_HUMAN STANDARD; PRT; 416 AA.
AC Q9UI08; Q95884; Q8TBV1; Q9UF25; Q9UIC2;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Ena/vasodilator stimulated phosphoprotein-like protein (Ena/VASP-like
DE protein).
GN Name=EVL; Synonyms=RNBE;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 2).
RA Ohta S., Mineta T., Kimoto M., Tabuchi K.;
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 1).
RA TISSUE=Adrenal gland;
RX MEDLINE=20402571; PubMed=10931946; DOI=10.1073/pnas.1602709997;
RA Hu R.-M., Han Z.-G., Song H.-D., Peng Y.-D., Huang Q.-H., Ren S.-X.,
RA Gu Y.-J., Huang C.-H., Li Y.-B., Jiang C.-L., Fu G., Zhang Q.-H.,
RA Gu B.-W., Dai M., Mao Y.-F., Gao G.-F., Rong R., Ye M., Zhou J.,
RA Xu S.-H., Gu J., Shi Y.-X., Jin W.-R., Zhang C.-K., Wu T.-M.,
RA Huang G.-Y., Chen Z., Chen M.-D., Chen J.-L.;
RT "Gene expression profiling in the human hypothalamus-pituitary-adrenal
RT axis and full-length cDNA cloning".
RL Proc. Natl. Acad. Sci. U.S.A. 97:9543-9548(2000).
RN [3]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORMS 1 AND 2).
RC TISSUE=Lung, and Pancreas;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Maman A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalley D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Maria M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences".
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 55-416.
RC TISSUE=Brain;
RA Mei G., Yu W., Gibbs R.A.;

Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
[5]
RN NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 163-364.
RP TISSUE=Uterus;
RG The German cDNA consortium;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
[6]
RN INTERACTION WITH SEMA6A.
RP PubMed=10993894; DOI=10.1074/jbc.M006316200;
RA Klostermann A., Lutz B., Gertler F., Behl C.;
RT "The orthologous human and murine semaphorin 6A-1 proteins (SEMA6A-1/sem6a-1) bind to the enabled/vasodilator-stimulated phosphoprotein-like protein (EVL) via a novel carboxyl-terminal zyxin-like domain.";
RL J. Biol. Chem. 275:39647-39653 (2000).
[7]
RN PHOSPHORYLATION SITES SER-329 AND SER-331.
RP PubMed=15302935; DOI=10.1073/pnas.0404720101;
RA Beausoleil S.A., Jedrychowski M., Schwartz D., Elias J.E., Villen J., Li J., Cohn M.A., Cantley L.C., Gygi S.P.;
RT "Large-scale characterization of HeLa cell nuclear phosphoproteins.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:12130-12135 (2004).
CC -|- FUNCTION: Enhances actin nucleation and polymerization (By similarity).
CC -|- SUBUNIT: Binds to the SH3 domains of ABL1, LYN and SRC. Also binds to profilin and the WW domain of APBB1/FE65 (By similarity). Binds to SEMA6A.
CC -|- INTERACTION: P08631:HCK; NbExpel; IntAct=EBI-346653, EBI-346340;
CC -|- SUBCELLULAR LOCATION: Focal adhesions (By similarity).
CC -|- ALTERNATIVE PRODUCTS:
CC -|- Event=Alternative splicing; Named isoforms=2;
CC Name=2;
CC IsoId=Q9UI08-1; Sequence=Displayed;
CC Name=1;
CC IsoId=Q9UI08-2; Sequence=VSP_004044;
CC SH3 domains of ABL and SRC (By similarity).
CC -|- SIMILARITY: Contains 1 WH1 domain.
CC This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.

DR EMBL; AF052504; AAP21709.1; -; mRNA.
DR EMBL; AF112209; AAP17197.1; -; mRNA.
DR EMBL; BC023997; AAH23997.1; -; mRNA.
DR EMBL; BC023358; AAH32358.1; -; mRNA.
DR EMBL; AF131766; AAD20040.1; -; mRNA.
DR EMBL; AL133642; CAB63763.2; -; mRNA.
DR HSSP; P70429; LOC6.
DR SMR; P70429; LOC6.
DR IntAct; Q9UI08; -.
DR EMBL; ENSG00000196405; Homo sapiens.
DR HGNC; HGNC:20234; EVL.
DR H-InvDB; HIX0011964; -.
DR GO; GO:005737; C:cytoplasm; ISS.
DR GO; GO:0005225; C:focal adhesion; ISS.
DR GO; GO:0030027; C:lamellipodium; ISS.
DR GO; GO:0005522; F:profilin binding; ISS.
DR GO; GO:0017124; F:SH3 domain binding; ISS.
DR GO; GO:0007015; P:actin filament organization; TAS.
DR GO; GO:0030048; P:actin filament-based movement; ISS.
DR GO; GO:0045010; P:actin nucleation; ISS.
DR GO; GO:0008154; P:actin polymerization and/or depolymerization; ISS.
DR GO; GO:0007411; P:axon guidance; ISS.
DR GO; GO:0051016; P:barbed-end actin filament capping; ISS.
DR GO; GO:0007166; P:cell surface receptor linked signal transdu. . .; NAS.
DR GO; GO:0007399; P:neurogenesis; NAS.
DR GO; GO:0030168; P:platelet activation; ISS.
DR InterPro; IPR000697; EVH1.
DR InterPro; IPR011993; PH_type.

DR Pfam; PF00568; WH1; 1.
DR SMART; SMO0461; WH1; 1.
DR PROSITE; PS00229; WH1; 1.
KW Alternative splicing; Phosphorylation; SH3-binding.
FT DOMAIN 1 112 WH1.
FT COMPIAS 162 206 Pro-rich.
FT MOD_RES 329 331 Phosphoserine.
FT MOD_RES 331 331 Phosphoserine.
FT VARSPPLIC 1 1 M -> MAT (in isoform 1).
FT /FTId=VSP_004044.
FT CONFLICT 201 201 P -> S (in Ref. 2).
FT CONFLICT 364 364 M -> Y (in Ref. 5).
SQ SEQUENCE 416 AA; 44620 MW; AD5B67458755D659 CRC64;
Query Match 83.6%; Score 56; DB 1; Length 416;
Best Local Similarity 80.0%; Pred. No. 1.5e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Qy 1 PPPYPPPPIP 10
Db 182 PPPPPPPPPV 191
RESULT 44
Q7Z522 HUMAN
ID Q7Z522 HUMAN PRELIMINARY; PRT; 418 AA.
AC Q7Z522;
DT 01-OCT-2003 (T-EMBLrel. 25, Created)
DT 01-OCT-2003 (T-EMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
DS B6.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Wan Y.Z., Yu L., Yue P., Tu Q., Fu S.N., Zhao S.Y.;
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
[2]
RP NUCLEOTIDE SEQUENCE.
RA Zhang M., Yu L., Zhou Y., Hu P.R., Xin Y.R., Zhao S.Y.;
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF087843; AAP97156.1; -; mRNA.
DR HSSP; P50552; IEGX.
DR SMR; Q7Z522; 4-115.
DR EMBL; ENSG00000196405; Homo sapiens.
DR InterPro; IPR000697; EVH1.
DR InterPro; IPR011993; PH_type.
DR InterPro; IPR001960; WH1.
DR Pfam; PF00568; WH1; 1.
DR SMART; SMO0461; WH1; 1.
SQ SEQUENCE 418 AA; 44819 MW; 02D02741E5EE7219 CRC64;
Query Match 83.6%; Score 56; DB 2; Length 418;
Best Local Similarity 80.0%; Pred. No. 1.5e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Qy 1 PPPYPPPPIP 10
Db 184 PPPPPPPPPV 193
RESULT 45
Q5R896 PONPY
ID Q5R896 PONPY PRELIMINARY; PRT; 422 AA.
AC Q5R896;
DT 01-FEB-2005 (T-EMBLrel. 29, Created)
DT 01-FEB-2005 (T-EMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (T-EMBLrel. 29, Last annotation update)
DE Hypothetical protein DKFZp468J0419.
GN Name=DKFZp468J0419;

```

OS Pongo pygmaeus (Orangutan).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Pongo.
OX NCBI_TaxID=9600;
RN NUCLEOTIDE SEQUENCE.
RC TISSUE=Heart;
RG The German cDNA Consortium;
RA Ottenwaelder B., Obermaier B., Deutschenbaur S., Schaiipp A.,
RA Mewes H.W., Weil B., Anlid C., Osanger A., Fobo G., Han M., Wiemann S.;
RL Submitted (NOV-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR859858; CAH92014.1; -; mRNA.
DR SRR; OSR896; 8-119.
DR InterPro; IPR000697; EVH1.
DR InterPro; IPR001960; WH1.
DR Pfam; PF00568; WH1; 1.
DR SMART; SM00461; WH1; 1.
DR KW Hypothetical protein.
SQ SEQUENCE 422 AA; 45314 MW; ACB6E9F6CBA77D99 CRC64;

Query Match 83.6%; Score 56; DB 2; Length 422;
Best Local Similarity 80.0%; Pred. NO. 1.5e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPYPPPPPIP 10
Db 188 PPPPPPPPPVP 197

RESULT 46
Q7KK41_DROME
ID Q7KK41_DROME PRELIMINARY; PRT; 437 AA.
AC Q7KK41;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE SNF1A.
GN Name=SNF1A;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN NUCLEOTIDE SEQUENCE.
RC [1]
RA Rubin G.M., Wan K.H., Harvey D., Lewis S.E., Brokstein P., Tsang G.,
RA Agbayani A., Arcaina T., Baxter E., Blazej R.G., Butenhoff C.,
RA Champe M., Chavez C., Chew M., Doyle C.M., Farfan D.E., Frise E.,
RA Galle R., George R.A., Harris N.L., Hoskins R.A., Evans-Holm M.,
RA Houston K.A., Hummaati S.R., Kim E., Li P., Moshrefi M., Pacleb J.M.,
RA Park S., Sequeira A., Sethi H., Snir E., Svirskas R.R., Weinburg T.,
RA Celniker S.E.;
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
DR EMBL; AF181649; AAD55435.1; -; mRNA.
DR GO; GO:000524; F-ATP binding; IEA.
DR GO; GO:0004674; F-protein serine/threonine kinase activity; IEA.
DR GO; GO:0006468; P-protein amino acid phosphorylation; IEA.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR008271; Ser_thr_kin AS.
DR InterPro; IPR002290; Ser_thr_kinase.
DR Pfam; PF00069; Pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR ATP-binding; Kinase; Nucleotide-binding;
DR KW Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 437 AA; 48302 MW; A5CACB2A0473B288 CRC64;

Query Match 83.6%; Score 56; DB 2; Length 437;
Best Local Similarity 80.0%; Pred. NO. 1.5e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPYPPPPPIP 10
Db 188 PPPPPPPPPVP 197

RESULT 47
Q6YZ63_ORYSA
ID Q6YZ63_ORYSA PRELIMINARY; PRT; 439 AA.
AC Q6YZ63;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein OSJNB0013K01.33.
GN Name=OSJNB0013K01.33;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN NUCLEOTIDE SEQUENCE.
RC [1]
RA Sasaki T., Matsumoto T., Katayose Y.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 2, BAC
clone:OSJNB0013K01."
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP005538; BADI7380.1; -; Genomic_DNA.
DR Gramene; Q6YZ63; -.
KW Hypothetical protein.
SQ SEQUENCE 439 AA; 46263 MW; A8070E962A857306 CRC64;

Query Match 83.6%; Score 56; DB 2; Length 439;
Best Local Similarity 80.0%; Pred. NO. 1.6e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPYPPPPPIP 10
Db 96 PPPPPPPPPMP 105

RESULT 48
Q4R8F0_MACFA
ID Q4R8F0_MACFA PRELIMINARY; PRT; 480 AA.
AC Q4R8F0;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Testis cDNA clone: QtsA-12592, similar to human hypothetical protein
LOC195977 (LOC195977).
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
OC Cercopithecoidea; Cercopithecinae; Macaca.
OX NCBI_TaxID=9541;
RN NUCLEOTIDE SEQUENCE.
RC [1]
RA International consortium for macaque cDNA sequencing, analysis;
RT "DNA sequences of macaque genes expressed in brain or testis and its
evolutionary implications."
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
RN NUCLEOTIDE SEQUENCE.
RC [2]
RA Osada N., Hirata M., Tanuma R., Kusuda J., Hida M., Suzuki Y.,
RA Sugano S., Gojobori T., Shen J.C.-K., Wu C.I., Hashimoto K.;
RT "Substitution rate and structural divergence of 5'UTR evolution:
Comparative analysis between human and cynomolgus monkey cDNAs."
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB168502; BA800622.1; -; mRNA.
KW Hypothetical protein.
SQ SEQUENCE 480 AA; 53172 MW; 5EC5C743D43CC1B1 CRC64;

Query Match 83.6%; Score 56; DB 2; Length 480;
Best Local Similarity 80.0%; Pred. NO. 1.5e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPYPPPPPIP 10
Db 220 PPPPPPPPPVP 229

```

Best Local Similarity 80.0%; Pred. No. 1.7e+02; DB 2; Length 557;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPYPPPPPIP 10
Db 251 PPPSPPPPPLP 260

RESULT 49

Q73XK9_MYCPA PRELIMINARY; PRT; 489 AA.
AC Q73XK9;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN OrderedLocustNames=MAP2300c;
OS Mycobacterium paratuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium avium complex (MAC).
OX NCBI_TaxID=1770;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=k10;
RA Li L., Bannantine J., Zhang Q., Amons A., Alt D., Kapur V.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB017235; AAS04617.1; -; Genomic_DNA.
DR InterPro; IPR004255; UPF0089.
DR Pfam; PF03007; UPF0089; 1.
KW Complete proteome.
SQ SEQUENCE 489 AA; 52083 MW; 40A98903DB6D6890 CRC64;

Query Match 83.6%; Score 56; DB 2; Length 489;
Best Local Similarity 80.0%; Pred. No. 1.7e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPYPPPPPIP 10
Db 167 PPPPSPPPPLP 176

RESULT 50

Q4R7B7_MACFA PRELIMINARY; PRT; 557 AA.
AC Q4R7B7;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Testis cDNA, clone: QtaA-15671, similar to human hypothetical protein LOC195977 (LOC195977).
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Cercopithecoidea; Cercopithecidae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA International consortium for macaque cDNA sequencing, analysis; "DNA sequences of macaque genes expressed in brain or testis and its evolutionary implications."
RT Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Osada N., Hirata M., Tanuma R., Kusuda J., Hida M., Suzuki Y., Sugano S., Gojobori T., Shen J.C.-K., Wu C.I., Hashimoto K.; "Substitution rate and structural divergence of 5'UTR evolution: Comparative analysis between human and cynomolgus monkey cDNAs.";
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
RW EMBL; AB168903; BAE01005.1; -; mRNA.
KW Hypothetical protein.
SQ SEQUENCE 557 AA; 62019 MW; B524252DBE18FAC CRC64;

Query Match 83.6%; Score 56; DB 2; Length 557;
Best Local Similarity 80.0%; Pred. No. 2e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Qy 1 PPPYPPPPPIP 10
Db 400 PPPSPPPPPLP 409

Search completed: April 6, 2006, 09:38:59
Job time : 172.895 secs